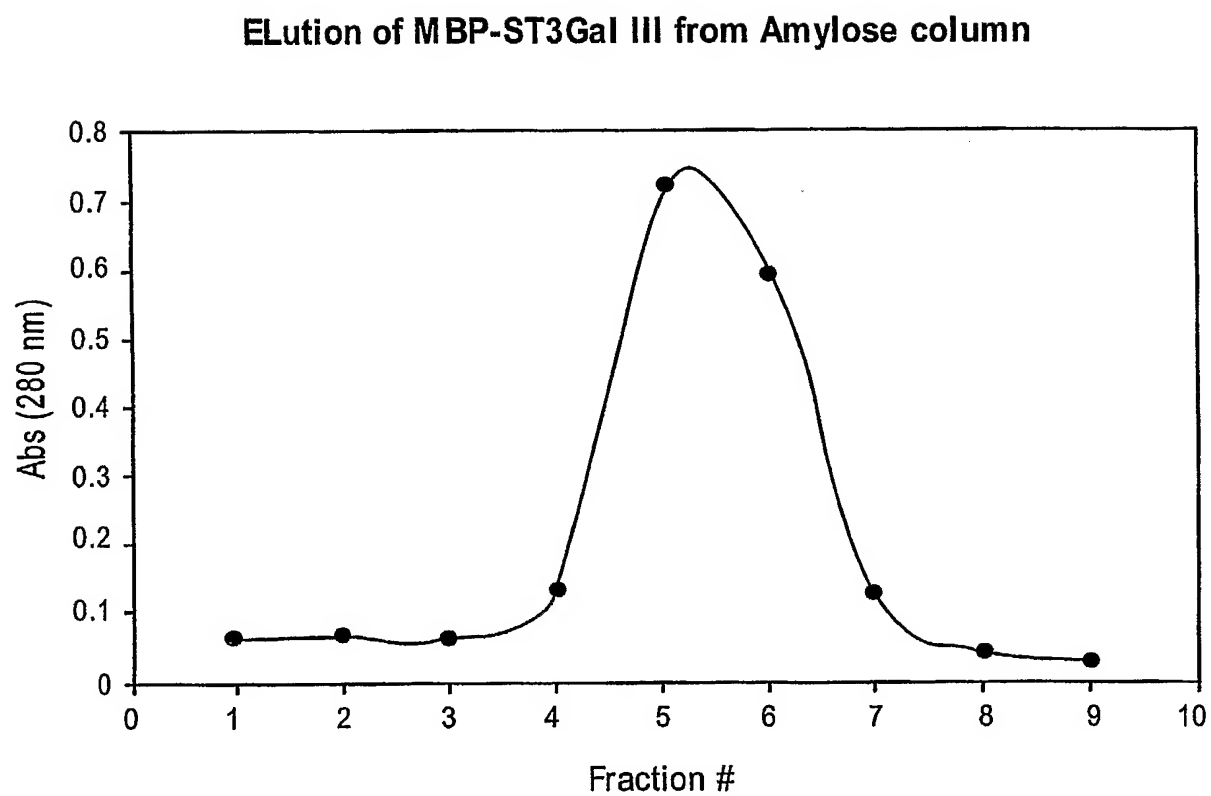


1/54

	1 mM	0.1 mM	0.3 mM	mM	mM	0.055%	550 mM	1.1 mM	2.2 mM	2.2 mM	CaCl ₂	440 mM	550 mM	Activity
#	GSH	GSSG	LM	NaCl	KCl	PEG 3350	GndHCl	EDTA	MgCl ₂			Sucrose	L- Arg	U/g IB
2 (55 mM MES pH 6.5)	+	+	+	10.56	0.44	0	+	0	+		+	0	0	0
3 (55 mM MES pH 6.5)	+	+	0	10.56	0.44	+	+	+	0		0	+	+	0
#5 (55 mM MES pH 6.5)	+	+	0	264	11	0	0	0	+		+	+	0	0
#8 (55 mM MES pH 6.5)	+	+	+	264	11	+	0	+	0		0	0	+	40.00
#10 (55 mM MES pH 8.2)	+	+	+	10.56	0.44	0	0	+	0		0	+	0	0
#11 (55 mM Tris pH 8.2)	+	+	0	10.56	0.44	+	0	0	+		+	0	+	105.26
#13 (55 mM Tris pH 8.2)	+	+	0	264	11	0	+	+	0		0	0	0	15.65
#16 (55 mM Tris pH 8.2)	+	+	+	264	11	+	+	0	+		+	+	+	48.70

FIG. 1

2/54

**FIG. 2**

3/54

**ST3 Gal III activities of the Amylose purified
refolded MBP-ST3Gal III fractions**

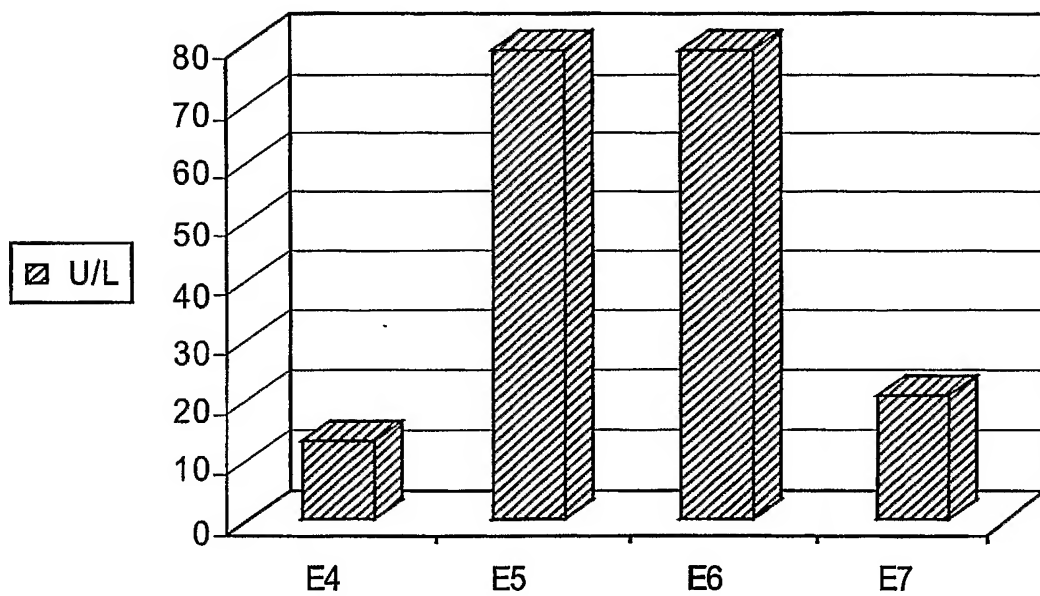
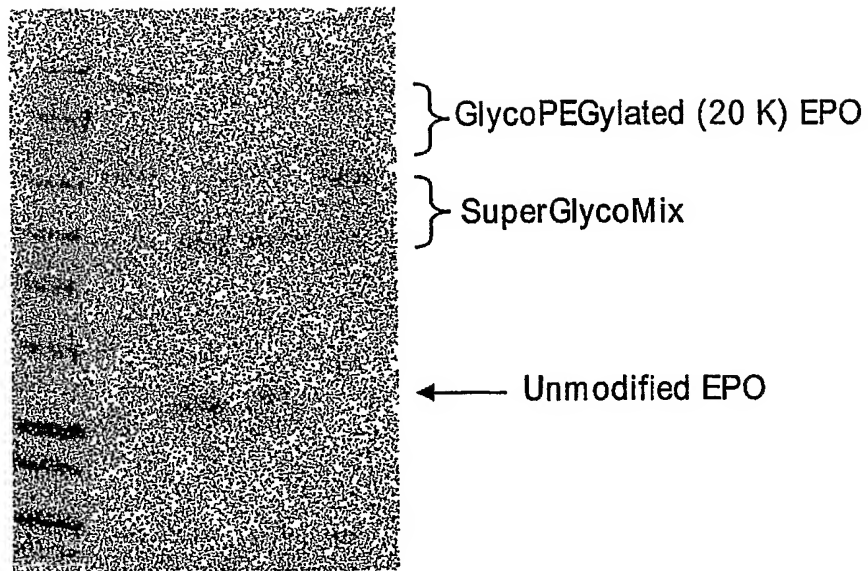


FIG. 3

4/54

**FIG. 4**

5/54



GlycoPEGylation (20 K) of EPO

FIG. 5

10 20 30 40 50 60
/usr/t MLKKQSAGLV LWGAILF VAWNAL LLLFFWTR PAPGRPPSVSALDGD PASLTRE VIRLAQD
::: :::
P27115 MLKKQSAGLV LWGAILF VAWNAL LLLFFWTR FVPSRL PSDNAL DDDPASLTRE VIRLAQD
10 20 30 40 50 60

70 80 90 100 110
/usr/t AEVELERQRGLLQQIGD--ALSSQRGRVP TAAAPPAQPRVPVTPAPAVIPILVIACDRSTV
::: :
P27115 AEVELERQRGLLQQIREH HALWSQRWKVPTAAPPAQPHVFPVTPPPAVIPILVIACDRSTV
70 80 90 100 110 120

120 130 140 150 160 170
/usr/t RRCLDKLLHYRPSAELEFPIIVSQDCGHEETAQA IASYGS AVTHIRQPDLSSI AVPPDHRK
::: :::
P27115 RRCLDKLLHYRPSAELEFPIIVSQDCGHEETAQVI ASYGS AVTHIRQPDL SNIAVQPDHRK
130 140 150 160 170 180

180 190 200 210 220 230
/usr/t FQGY YKIARHYRWALGQVERQFRFPAAVVVEDD LEVAPDFFEYFRATYP LLKADPSLWCV
::: :::
P27115 FQGY YKIARHYRWALGQIFHNFNYPAAVVVEDD LEVAPDFFEYFOATYP LLKADPSLWCV
190 200 210 220 230 240

240 250 260 270 280 290
/usr/t SAWNDNGKEQMVDASRP ELLYRTDFFPGLGWILL LAELWA ELEPKWP KAFWD DMRRPEQR
::: :::
P27115 SAWNDNGKEQMVDSSKP ELLYRTDFFPGLGWILL LAELWA ELEPKWP KAFWD DMRRPEQR
250 260 270 280 290 300

300 310 320 330 340 350
/usr/t QGRACIRPEISRITMIFGRKGVS HGOFFDQHLKF IKLNQQFVHF TQLDLSYLQ REAYDRDF
::: :::
P27115 KGRACVRPEISRITMIFGRKGVS HGOFFDQHLKF IKLNQQFVPFTQLDLSYLQ QEAYDRDF
310 320 330 340 350 360

360 370 380 390 400 410
/usr/t IARVYGAPQLQVEKV RTNDRKELGEVRVQYTGRDSFK AFAKALGVMD DLKSGVPRAGYRG
::: :::
P27115 IARVYGAPQLQVEKV RTNDRKELGEVRVQYTGRDSFK AFAKALGVMD DLKSGVPRAGYRG
370 380 390 400 410 420

420 430 440
/usr/t IVTFQFRGRRVHLAPP PTWEGYDPSWN
::: :::
P27115 IVTFLFRGRRVHLAPP QTWDGYDPSWT
430 440

FIG. 6

FIG. 6

GnT1 Cys121Ser mutant

avipilviacdrstvrirldklhlyrpsaelpiivsqdcgheetaqaiaisygsavthirqpdlssiaavppdhrkfqggyykiarhyrwa
lgqvfrqfrfpaavvveddleavpddfeyfratypilkadpslwcvswndngkeqmvdasrpellyrtdffpglgwllaelwae
lepkwpkafwddwmirpeqrqgracirpeisrmtftgrkgvshgqffidqhlkfiklnqfvhftqlidsylqreaydrdflarvyg
apqlqvekvtrndrkelgevrvqytrgrdsfkafakalgvmddksgvpragygivtfqfpgrryhlappptwewgydpwn*

Gcgggtgattcccatcctgtgtcatcgcctgtgacgcagcactgttcggcggcgtctcttagacaagctgtgcattatcgccctcgggctga
gctcttcccatcatcgttagccaggactgcgggcacgagagagacggcccgaggccatcgctcctacggcagcgcggtcacgcaca
tcgggcagcccgacctgagcagcattgggtgcgcggcagcagcaccgcaagttccagggtctactacaagatcggcggccactacccg
ctggggcgtcgggcaggtcttcggcagtttcgcttcccgccggcggcgtggtggtgaggtgacacctggaggtggccccggacttctt
cgagtagcttccgggcccactatccgctgctgaaggccgacccctccctgtgtgctcgcctgggaatgacaacggcgaaggagca
gatggtggagcgcagcagcctgagctgtctctacccgacccgacttttccctggcctgggctgtgtgcccggagctctgggct
gagctggagcccaagtggccaaaggcccttctgggacgacgtggtgacggcggcggcggcagcagcgaggcgccgtgcatacg
ccctgagatctcaagaacgatgacaccttggccgcagagggtgtgagccacggcgagttcttgaccagcacctcaagttaacagctga
accagcagttgtgtgcacttcacccagcgtggacgtgtcttacctgcagcgggagggcctatgacccgagatttccgcggcgtctacgggt
gctccccagctgcaggtggagaaagtgaggaaccaatgacgcggaaggagctggggggaggtgcgggtgcagtatacggggcaggga
cagcttcaaggcttcccaaggctctgggtgtcatggatgaccttaagtcgggggggttcggagagctgggtaccgggggtattgtcacctt
ccagttcccgggccgctgtgtccacctggcgcccccacgcagtgaggggagggctatgtatctctctagctgggaattag

7/54

FIG. 7

GnT1 Cys121Asp

avipilviacdrstvrrdldkllhyrpsaelfpriivsqdcgheetaqaiasygsavthirqpdlssiavpdpdhrkfqggykiahryrwa
 lgqvfrqfrpaavvveddleavpddfeyfratypllkadpslwcvsawndngkeqmvdasrpellyrtdffpglglwllaelwae
 lepkwpkafwddwmurpeqrqgracirpeisrtmtfgrkgvshgqffdqhlkfiklnqfwhftqldlsylqreaydrdrflarvyg
 apqlqvekvtrndrkelgevrvqy'tgrdsfkafakalgvmddllksgvpragrygivtfqfpgrrvhlappptwegydpswn*

Gcgggtgattcccatcctggtcatcgccctgtgaccgcagcactgttcggcggcgtatctagacaagctgctgcattatcgccctcggcctg
 agctcttcccatcatcgtagccaggactgcggggcacgagagagacggccagggccatcgctctctacggcagcgcggtcacggcac
 atccggcagcggccgacctgagcagcattggcggcgccggacacccgcaagttccaggggctactacaagatcgcgcgccactacc
 gctggggcgctggggccagggtcttccggcaggtttcgtctcccgcgcgccggtggggagggatgacctggggagggcgcccgggacttct
 tcgagtagtcttcggggccacctatccgctgctgaaaggccgacccctccctggtgggtgctctggccctgggaatgacaaaggcgaaggagc
 agatgggtggagcgccagcgagcgtgagctgctctacggcaccgacttttccctggccctggcgctggcgctggctggcgctggcgctggcg
 tggagctgggagggcccaagggcccttctggggagcagactggatggcgcgcgccggcgagcagcgcgcgcgcgcgcgcgcgcgcgcgcg
 gcccctgagagatcacaagaaacgatgacctttggccgcaagggtggagccacggggcaggttcttgaccagcaoctcaagttatcaagctg
 aaaccagcaggtttggtgacgttcaaccagctggagacctggttacctggcagcggggagggccctatggaccgagatttctcgcccgctctacgg
 tggctcccaagcgtgcaggtggagaaagtgaagaccaatggacccggaaaggagctggggggaggggtggcgaggtatatacgggcgagggga
 caggcttcaagggttccgcaaggctctgggggtgctcatggatgaccttaagctcgggggggttccggagagctggcgctaccgggggtattgtcacctt
 ccaggttcccgggcgccgctgtgtcccaacctggcgcccccacacgacgtggggaggggtatgatcctagctgggaattag

FIG. 8

GnT1 Cys121Thr

avipilviacdrtvrrtdkllhyrpsaelpiivsqdcgheetaqaiasygsavthirqpdlssiaivppdhrkfqgvykiarhyrwal
gqvfrqfrpaa vvveddle vapdfeyfratypllk adpslwcv sawndngkeqmv dasrpellyrtddfglgwllael wael
epkwpkafwdwmrtp eqqgracirpeistrmtfgrkgvshgqffdlhlfiklnqqf vhfhtql dlsylqreaydrdrflarvyg
apqlqavekvrtnrkelgevrvqytrdsfkafakalgvmdllksgvpragyrvitvtfqgrrvhlappptwgydpswn*

[illegible]

Fig. 6

GnT1 Cys121Ala

avipilviacdrstvrraldklhlyrpsaelfpriivsqdcgheetaqaiasygsavthirqpdlssiavppdhrkfqggykiarhyrwa
 lgqvfrqfrfpaavvveddleavpddfeyfratypllkadpslwcvsawndngkeqmvdasrpellyrtdffplglgwllaelwae
 lepkwpkafwddwmrpeqrgracirpeisrntnftgrkgvshgqffdqhlkfiklnqfwhftqldsylqreaydrdrflarvyg
 apqlvkvtrndrkelgevrvqytrgrdsfkafakalgvmdldksvpragrygvtqfpgfrrvhlappptwgydpsswn*

Gcgggtattccatcccttggtcatcgccctgtgacccgcagcactgttcggcggcccttagacaagctgctgcattatcgccctcggtcg
 agctcttcccatcatcggttagccaggaactgcggggcacgagggagacggccagggccatcgctctctacggcgagcgcggtcacggcac
 atccggcagccccgacctgagcagcattgcgggtgcccggaccacggcaagtccagggctactacaagatcgcgcgccactacc
 gctggggcgctggggccaggtcttccggcaggttctgcttccccggcggcggctggtggggagggatgacctgggaggtggccccgggacttct
 tcgagttactttcggggccaccatccgctgctgaaaggccgagccccctccctgtgggtgctctggcctgggaatgacaacgggcaaggagc
 agatgggtggagggccagcaggcctgagctgctctacccagcactttttccctgggctgggctgggctggtgtgtggccggagctctggggc
 tggagctgggagcccaagtgccaaaggcccttctggggacgactggatggcggcggccgggagcagcgggcgaggcgccctggcatac
 gcccctggagatctcaagaacgatgacctttggccgcaagggtgtgagccacggggcagttcttggaccagcaacctcaagttaacaaagctg
 aaccagcaggttgtgcacttcaaccagctggagctggtgacctgttaccttgacgggagggccctatgacccagagatttctcgcccgctctacggg
 tgcctcccagcagctgcaagggtggagaaagtggagaccaatgacccggaaaggagctggggggaggggtgcgggtgcaggtatatacggggcaggsga
 cagcttcaaggcttccgccaaggctctgggtgtcatggatgaccttaagtcgggggttccggagagctggggtaccgggggtattgtcacctt
 ccagttccccggggccgctgtgtccacacctggcggccccccacccagcgtgggaggggtatgatactatgctgggaattag

FIG. 10

12/54

Rat Liver ST3Gal III amino acid sequence:

MGLLVFVRNLLALCLFELVGLFYSAWKLHLLQWEDSNLSILSLDSAGQTLGTEYDRL
 GFLLKLD SKLP AELATKYANFSEGACKPGYASAMMTAIFPRFSKPAPMFLDDSF~~R~~KW
 ARIRFVPPFGIKGQDNLIKAILSVTKEYRLTPALDSLHCRRCIIVGNGGVLANKSLGS
 RIDDYDIVIRLNSAPVKGFEKDVGSKTTLRITYPEGAMQRPQYERDSLFVLAGFKW
 QDFKWLKYIVYKERVASDGFVKSVATRVPKPEPEIRILNPYFIQEAFTLIGLPFNN
 GLMGRGNIPTLGSAVVTMALDGCDEVAVAGFGYDMNTPNAPLHYHYYETVRMAAIKE
 SWTHNIQREKEFLRKLVKARVITDLSSGI

FIG. 12

14/54

Gln	Arg	Arg	Ser	Arg	Gln	Gly	Asn	Pro	Val	Ala	Pro	Ile	Lys	Thr	Pro
290						295					300				
Met	Ile	Ala	Gly	Gly	Leu	Phe	Val	Met	Asp	Lys	Phe	Tyr	Phe	Glu	Glu
305					310					315					320
Leu	Gly	Lys	Tyr	Asp	Met	Met	Met	Asp	Val	Trp	Gly	Gly	Glu	Asn	Leu
				325					330					335	
Glu	Ile	Ser	Phe	Arg	Val	Trp	Gln	Cys	Gly	Gly	Ser	Leu	Glu	Ile	Ile
			340					345					350		
Pro	Cys	Ser	Arg	Val	Gly	His	Val	Phe	Arg	Lys	Gln	His	Pro	Tyr	Thr
		355					360					365			
Phe	Pro	Gly	Gly	Ser	Gly	Thr	Val	Phe	Ala	Arg	Asn	Thr	Arg	Arg	Ala
	370					375					380				
Ala	Glu	Val	Trp	Met	Asp	Glu	Tyr	Lys	Asn	Phe	Tyr	Tyr	Ala	Ala	Val
385					390					395					400
Pro	Ser	Ala	Arg	Asn	Val	Pro	Tyr	Gly	Asn	Ile	Gln	Ser	Arg	Leu	Glu
				405					410					415	
Leu	Arg	Lys	Lys	Leu	Ser	Cys	Lys	Pro	Phe	Lys	Trp	Tyr	Leu	Glu	Asn
			420					425					430		
Val	Tyr	Pro	Glu	Leu	Arg	Val	Pro	Asp	His	Gln	Asp	Ile	Ala	Phe	Gly
		435					440					445			
Ala	Leu	Gln	Gln	Gly	Thr	Asn	Cys	Leu	Asp	Thr	Leu	Gly	His	Phe	Ala
	450					455					460				
Asp	Gly	Val	Val	Gly	Val	Tyr	Glu	Cys	His	Asn	Ala	Gly	Gly	Asn	Gln
465					470					475					480
Glu	Trp	Ala	Leu	Thr	Lys	Glu	Lys	Ser	Val	Lys	His	Met	Asp	Leu	Cys
				485					490					495	
Leu	Thr	Val	Val	Asp	Arg	Ala	Pro	Gly	Ser	Leu	Ile	Lys	Leu	Gln	Gly
			500					505					510		
Cys	Arg	Glu	Asn	Asp	Ser	Arg	Gln	Lys	Trp	Glu	Gln	Ile	Glu	Gly	Asn
		515					520					525			
Ser	Lys	Leu	Arg	His	Val	Gly	Ser	Asn	Leu	Cys	Leu	Asp	Ser	Arg	Thr
	530					535					540				
Ala	Lys	Ser	Gly	Gly	Leu	Ser	Val	Glu	Val	Cys	Gly	Pro	Ala	Leu	Ser
545					550					555					560
Gln	Gln	Trp	Lys	Phe	Thr	Leu	Asn	Leu	Gln	Gln					
				565					570						

FIG. 13A (CONT.)
 SUBSTITUTE SHEET (RULE 26)

15/54

Nucleic acid sequence

atgagggggc	gctogoggat	gctgctctgc	ttogocttcc	tgtgggtgct	gggcattogcc	60
tactacatgt	actcgggggg	cggctctgog	ctggcggggg	gogcgggggg	cggcggcggc	120
aggaaggagg	actggaatga	aattgacccc	attaaaaaga	aagaccttca	tcacagcaat	180
ggagaagaga	aagcacaag	catggagacc	ctccctccag	ggaaagtacg	gtggccagac	240
tttaaccagg	aagcttatgt	tggagggag	atggctccgt	cggggcagga	cccttaacgc	300
cgcaacaagt	tcaaccagg	ggagagtgat	aagcttogaa	tggacagagc	catccctgac	360
acccggcatg	accagtgtca	goggaagcag	tggcgggtgg	atctgcgggc	caaccagggtg	420
gtgatcacgt	ttcacaaatga	agccaggctg	gcccactcca	ggacogtggg	cagcgtgctt	480
aagaaaagcc	cgcgccatct	cataaaaaga	atcatcttgg	tggatgacta	cagcaatgat	540
octgaggag	gggctctctt	ggggaaaatt	gagaaaagtgc	gagttcttag	aaatgatoga	600
cgagaaggcc	tcattgogctc	acgggttogg	ggggccgatg	ctgcaccaagc	caaggctcgtg	660
accttctcgg	acagtcactg	cgagtgtaat	gagcactggc	tggagccct	octggaaaagg	720
gtggcggagg	acaggactcg	ggttgtgtca	cccatcatcg	atgtcattaa	tatggacaac	780
tttcagtatg	tgggggcac	tgctgacttg	aaggcggggt	ttgattggaa	cttgggtattc	840
aagtgggatt	acatgaogcc	tgagcagaga	aggtcccgcc	aggggaaccc	agtcgcccct	900
ataaaaaacc	ccatgattgc	tggctgggctg	tttgtgatgg	ataagttcta	ttttgaagaa	960
ctgggggaagt	acgacatgat	gatggatgtg	tggggaggag	agaacctaga	gatctogttc	1020
cgggtgtggc	agtgtggtgg	cagcctggag	atcatcccg	gcagcogtgt	gggacacgtg	1080
ttccggagc	agcaccocct	cacgttcccg	ggtggcagtg	gcactgtctt	tgcccgaaac	1140
acccgcgggg	cagcagaggt	ctggatggat	gaatacaaaa	atttctatta	tgcagcagtg	1200
octtctgcta	gaaacgttcc	ttatggaaat	attcagagca	gattggagct	taggaagaaa	1260
ctcagctgca	agoccttcca	atggtaoctt	gaaaatgtct	atocagagtt	aagggttcca	1320
gaccatcagg	atatagcttt	tggggccttg	cagcagggaa	ctaactgoc	cgacactttg	1380
ggacactttg	ctgatggtgt	ggttggagtt	tatgaatgtc	acaatgctgg	gggaaaccag	1440
gaatgggoc	tgacgaagga	gaagtogggt	aagcaccatg	atttgtgoc	tactgtggtg	1500
gaocgggcac	cgggctctct	tataaagctg	cagggctgoc	gagaaaatga	cagcagacag	1560
aaatgggaac	agatogaggg	caactccaag	ctgaggcacg	tgggcagcaa	octgtgoc	1620
gacagtogca	cggccaagag	cgggggccta	agcgtggagg	tgtgtggccc	ggccctttcg	1680
cagcagtgga	agttcacgct	caacctgcag	cag			1713

FIG. 13B
SUBSTITUTE SHEET (RULE 26)

16/54

**Δ51 UDP-N-acetylgalactosaminyltransferase 2, GalNAcT2,
nucleic acid and amino acid sequences**

Amino acid sequence

Lys	Lys	Lys	Asp	Leu	His	His	Ser	Asn	Gly	Glu	Glu	Lys	Ala	Gln	Ser
1				5					10					15	
Met	Glu	Thr	Leu	Pro	Pro	Gly	Lys	Val	Arg	Trp	Pro	Asp	Phe	Asn	Gln
			20					25					30		
Glu	Ala	Tyr	Val	Gly	Gly	Thr	Met	Val	Arg	Ser	Gly	Gln	Asp	Pro	Tyr
		35					40					45			
Ala	Arg	Asn	Lys	Phe	Asn	Gln	Val	Glu	Ser	Asp	Lys	Leu	Arg	Met	Asp
	50					55					60				
Arg	Ala	Ile	Pro	Asp	Thr	Arg	His	Asp	Gln	Cys	Gln	Arg	Lys	Gln	Trp
65					70					75					80
Arg	Val	Asp	Leu	Pro	Ala	Thr	Ser	Val	Val	Ile	Thr	Phe	His	Asn	Glu
				85					90					95	
Ala	Arg	Ser	Ala	Leu	Leu	Arg	Thr	Val	Val	Ser	Val	Leu	Lys	Lys	Ser
			100					105					110		
Pro	Pro	His	Leu	Ile	Lys	Glu	Ile	Ile	Leu	Val	Asp	Asp	Tyr	Ser	Asn
		115					120					125			
Asp	Pro	Glu	Asp	Gly	Ala	Leu	Leu	Gly	Lys	Ile	Glu	Lys	Val	Arg	Val
	130					135					140				
Leu	Arg	Asn	Asp	Arg	Arg	Glu	Gly	Leu	Met	Arg	Ser	Arg	Val	Arg	Gly
145					150					155					160
Ala	Asp	Ala	Ala	Gln	Ala	Lys	Val	Leu	Thr	Phe	Leu	Asp	Ser	His	Cys
				165					170					175	
Glu	Cys	Asn	Glu	His	Trp	Leu	Glu	Pro	Leu	Leu	Glu	Arg	Val	Ala	Glu
			180					185					190		
Asp	Arg	Thr	Arg	Val	Val	Ser	Pro	Ile	Ile	Asp	Val	Ile	Asn	Met	Asp
		195					200					205			
Asn	Phe	Gln	Tyr	Val	Gly	Ala	Ser	Ala	Asp	Leu	Lys	Gly	Gly	Phe	Asp
	210					215					220				
Trp	Asn	Leu	Val	Phe	Lys	Trp	Asp	Tyr	Met	Thr	Pro	Glu	Gln	Arg	Arg
225					230					235					240
Ser	Arg	Gln	Gly	Asn	Pro	Val	Ala	Pro	Ile	Lys	Thr	Pro	Met	Ile	Ala
				245					250					255	
Gly	Gly	Leu	Phe	Val	Met	Asp	Lys	Phe	Tyr	Phe	Glu	Glu	Leu	Gly	Lys
			260					265					270		
Tyr	Asp	Met	Met	Met	Asp	Val	Trp	Gly	Gly	Glu	Asn	Leu	Glu	Ile	Ser
		275					280					285			
Phe	Arg	Val	Trp	Gln	Cys	Gly	Gly	Ser	Leu	Glu	Ile	Ile	Pro	Cys	Ser

FIG. 14A
SUBSTITUTE SHEET (RULE 26)

17/54

290		295		300											
Arg 305	Val	Gly	His	Val	Phe 310	Arg	Lys	Gln	His	Pro 315	Tyr	Thr	Phe	Pro	Gly 320
Gly	Ser	Gly	Thr	Val 325	Phe	Ala	Arg	Asn	Thr 330	Arg	Arg	Ala	Ala	Glu 335	Val
Trp	Met	Asp	Glu 340	Tyr	Lys	Asn	Phe	Tyr 345	Tyr	Ala	Ala	Val	Pro 350	Ser	Ala
Arg	Asn	Val 355	Pro	Tyr	Gly	Asn	Ile 360	Gln	Ser	Arg	Leu	Glu 365	Leu	Arg	Lys
Lys	Leu 370	Ser	Cys	Lys	Pro	Phe 375	Lys	Trp	Tyr	Leu	Glu 380	Asn	Val	Tyr	Pro
Glu 385	Leu	Arg	Val	Pro	Asp 390	His	Gln	Asp	Ile	Ala 395	Phe	Gly	Ala	Leu	Gln 400
Gln	Gly	Thr	Asn	Cys 405	Leu	Asp	Thr	Leu	Gly 410	His	Phe	Ala	Asp	Gly 415	Val
Val	Gly	Val	Tyr 420	Glu	Cys	His	Asn	Ala 425	Gly	Gly	Asn	Gln	Glu 430	Trp	Ala
Leu	Thr	Lys 435	Glu	Lys	Ser	Val	Lys 440	His	Met	Asp	Leu	Cys 445	Leu	Thr	Val
Val	Asp 450	Arg	Ala	Pro	Gly	Ser 455	Leu	Ile	Lys	Leu	Gln 460	Gly	Cys	Arg	Glu
Asn 465	Asp	Ser	Arg	Gln	Lys 470	Trp	Glu	Gln	Ile	Glu 475	Gly	Asn	Ser	Lys	Leu 480
Arg	His	Val	Gly	Ser 485	Asn	Leu	Cys	Leu	Asp 490	Ser	Arg	Thr	Ala	Lys 495	Ser
Gly	Gly	Leu	Ser 500	Val	Glu	Val	Cys	Gly 505	Pro	Ala	Leu	Ser	Gln 510	Gln	Trp
Lys	Phe	Thr 515	Leu	Asn	Leu	Gln 520	Gln								

FIG. 14A (CONT.)

Nucleic acid sequence	18/54	
aaaaagaaag aoccttcatca cagcaatgga gaagagaaag cacaagcat ggagaoocac		60
octocagggg aagtaoogtg gacagacttt aaccaggaag cttatgttgg agggacgatg		120
gtocgctocg ggcaggacoc ttaoogocgc aacaagttca accaggtgga gagtataag		180
cttogaatgg acagagocat ooctgacac cggcatgaac agtgtcagcg gaagcagtg		240
oggggtgacac tgoogocac cagcgtggcg atcaogtttc acaatgaagc caggtogoc		300
ctactcagga oogtggtcag ogtgcttaag aaaagocgc ocatctcat aaaagaaac		360
atcttggtgg atgactacag caatgatoc gaggaogggg ctctcttggg gaaaattgag		420
aaagtgcag ttcttagaaa tgatogaag gaagocac tgoogtcag ggttogggg		480
gocgatgctg ocaagocaa ggtocgaac ttocggaca gtcactgga gtgtaatgag		540
cactggctgg agocococ tggaaagggg goggaagaca gactogggg tgtgtcaoc		600
atcatogatg tcattaatat ggacaacttt cagtatgtgg ggcacctgc tgactgaag		660
ggoggttttg attggaactt ggtattcaag tgggattaca tgacocga gcagagaag		720
toocggcagg ggaococagt ogococata aaaacocca tgattgctgg tgggctgtt		780
gtgatggata agttctatct tgaagaactg ggaagtaog acatgatgat ggatgtgtg		840
ggaggagaga aocagagat ctogttoooc gtgtggcag gtgggtggcag ootggagac		900
atocogtga gocgtgtggg acogtgttc oggaagcagc aococacac gttocoggg		960
ggcagtgga ctgtctttgc ogaacaac ogocggcag cagaggtctg gatggatga		1020
tacaaaaatt tctattatgc agcagtgoc tctgctagaa aogttoocta tggaaatatt		1080
cagagcagat tggagcttag gaagaaactc agctgaagc cttcaaag gtacottga		1140
aatgtctatc cagagttaag ggtocagac catcaggata tagcttttg ggocctgcag		1200
cagggaacta actgocoga cactttggga cactttgctg atgggtgtgt tggagttat		1260
gaatgtcaca atgctggggg aaocaggaa tgggocctga ogaaggagaa gtoggtgaag		1320
cacatggatt tgtgocctac tgtggtggac ogggcaocgg gctctcttat aaagctgcag		1380
ggctgcogag aaaatgacag cagacagaaa tgggaacaga togagggcaa ctocaagctg		1440
aggaogtgg gcagcaact gtgocggac agtogcaogg ocaagagogg gggocaaagc		1500
gtggaggtgt gtggocoggc octtogcag cagtgaagt tcaogtcaa octgcagcag		1560

FIG. 14B

SUBSTITUTE SHEET (RULE 26)

19/54

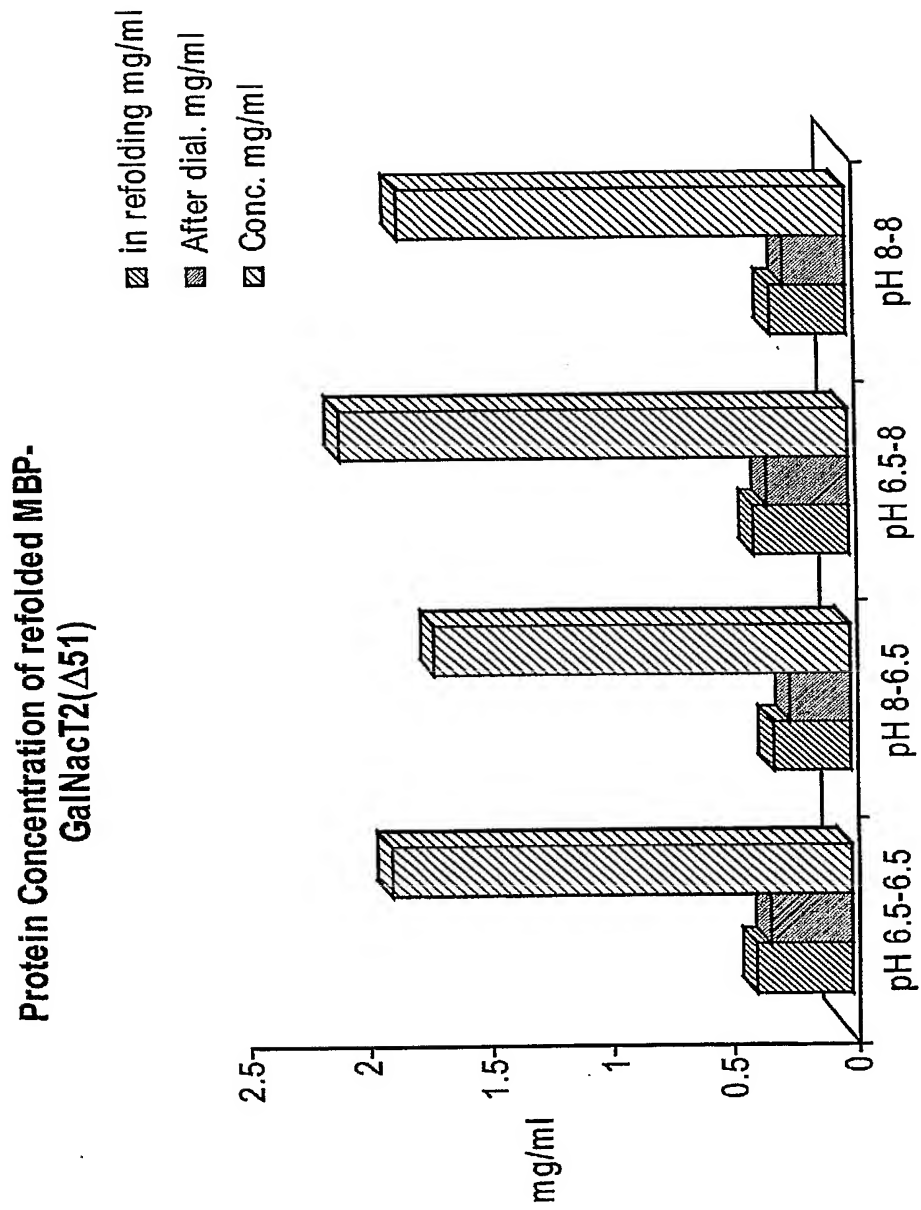


FIG. 15

20/54

GaINAct2 activities of refolded MBP-
GaINAct2(D51)

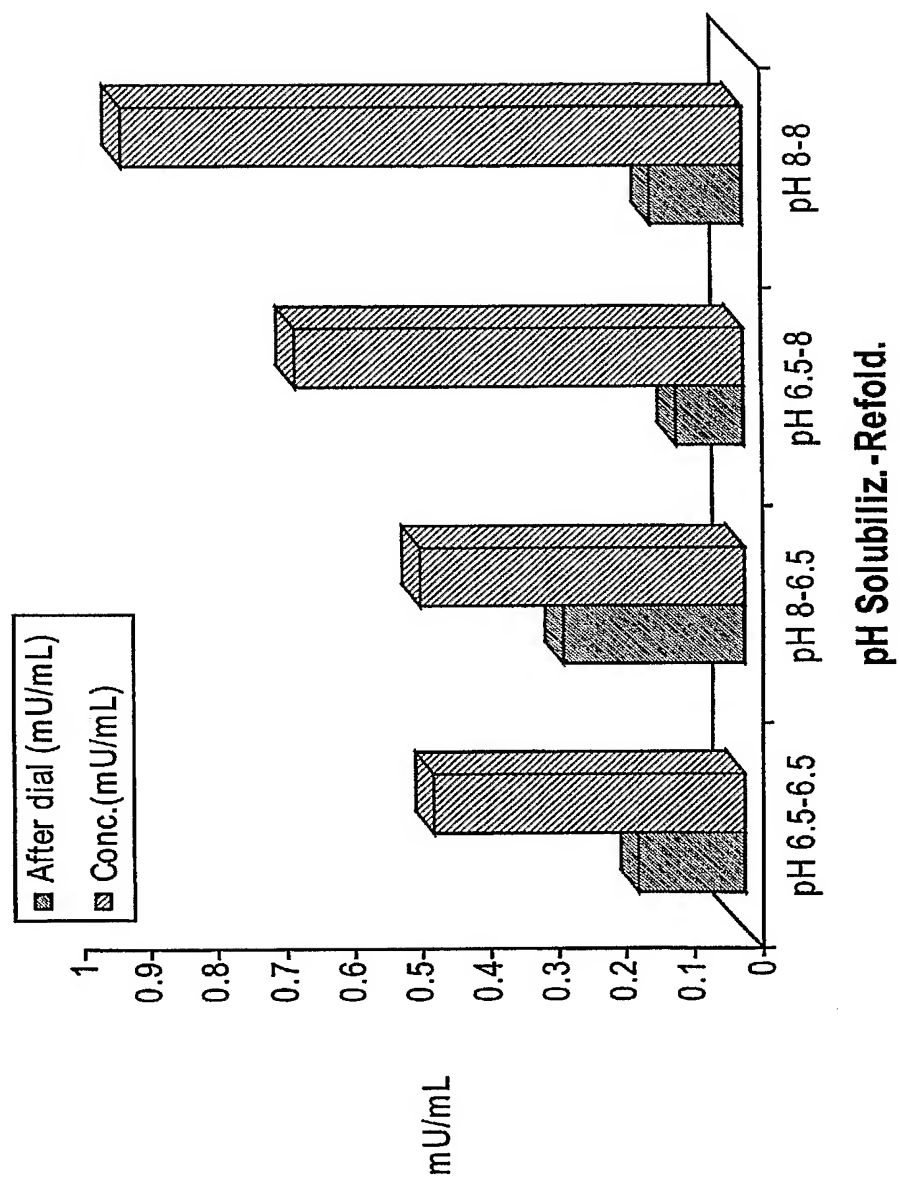


FIG. 16

pH Solubiliz.-Refold.

21/54

pH effect on the MBP-GalNAcT2(D51) specific activities

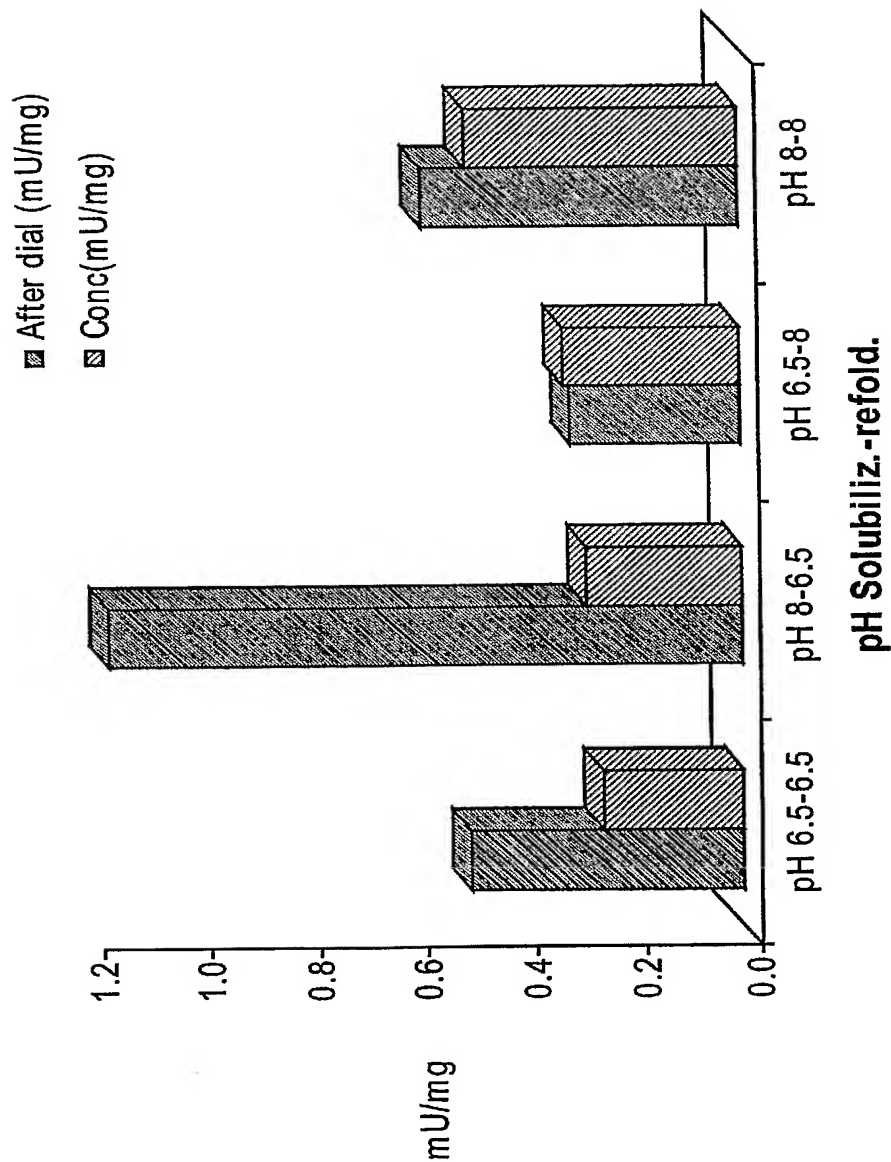


FIG. 17

22/54

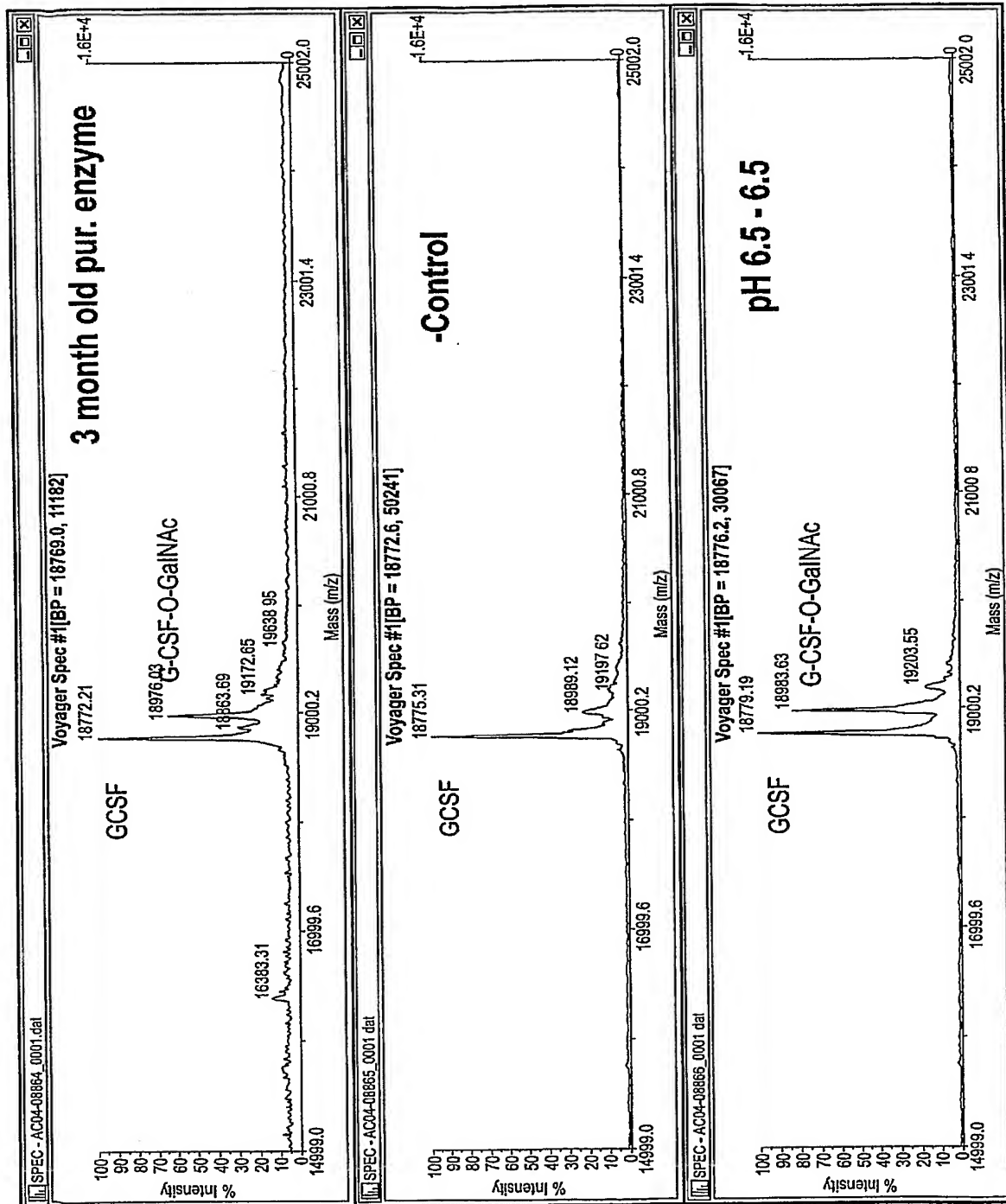


FIG. 18A

23/54

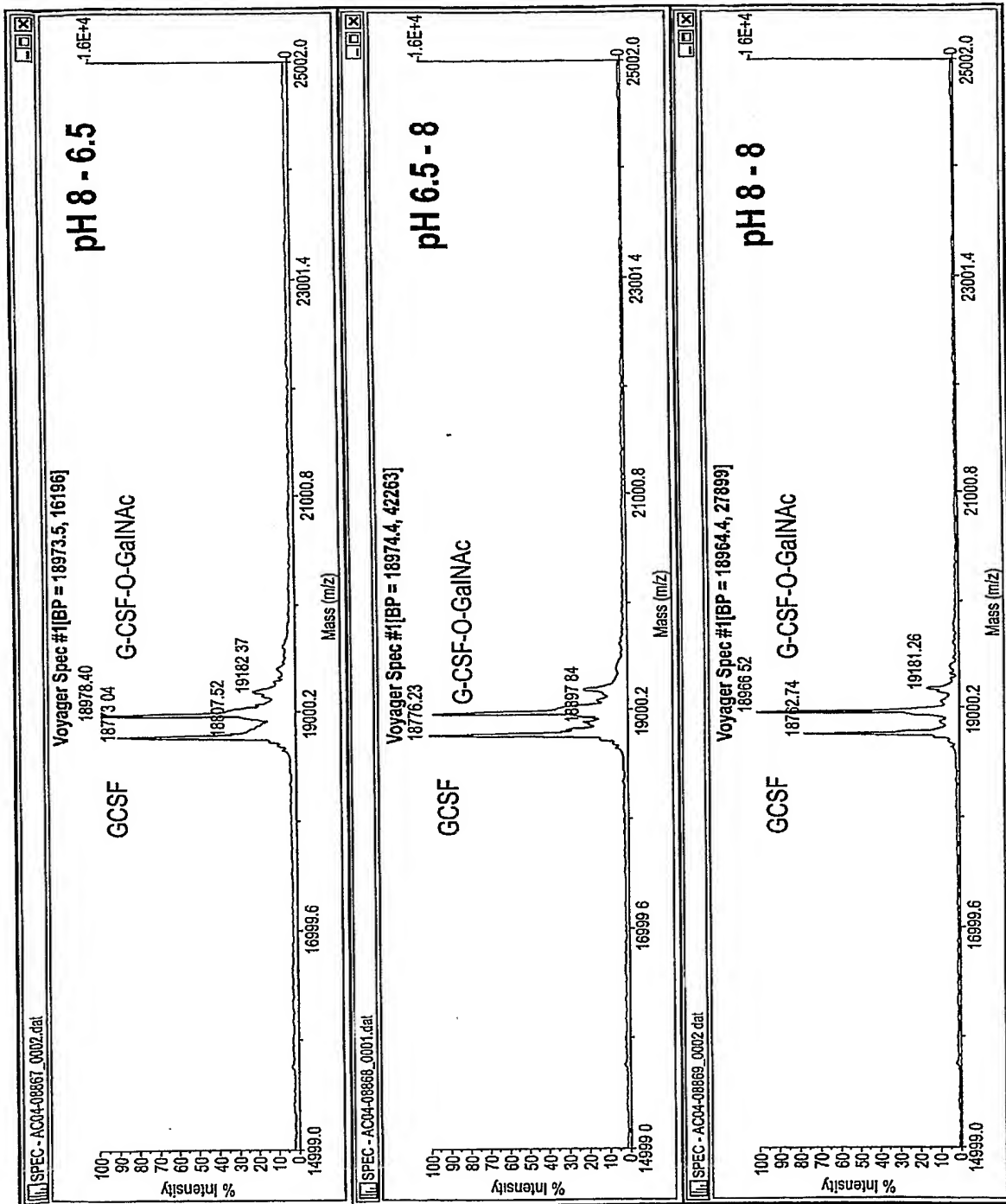


FIG. 18B

24/54

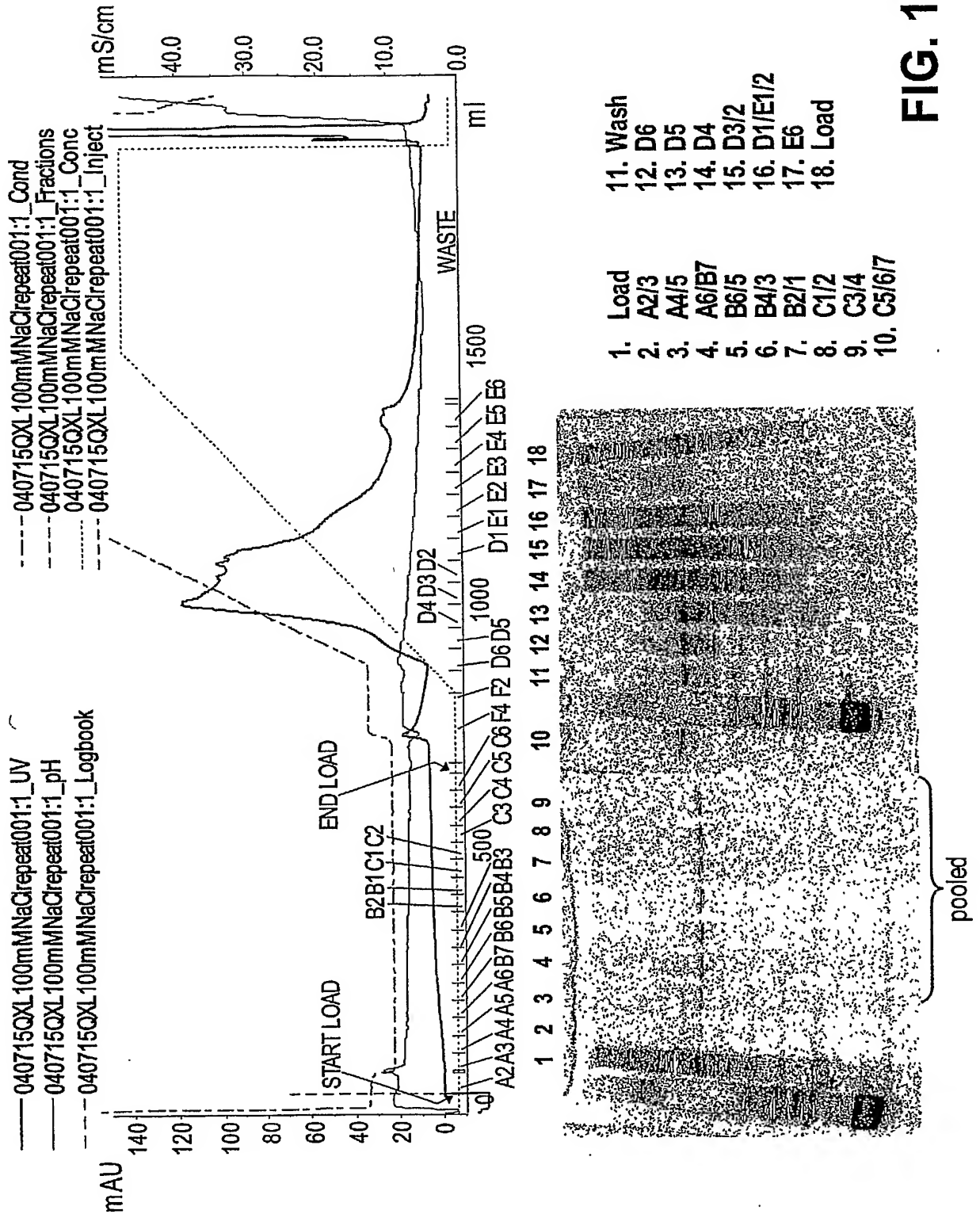


FIG. 19

25/54

	Volume (mL)	Activity (U/L)	A280
Load	890	1.5	0.110
FT _{A4-C4}	670	9.2	NA
FT _{C5-C7}	120	1.0	0
Wash	138	3.6	0.100
D6	45	4.5	0
D5	45	2.4	0.026
D4	45	2.0	0.108
D3/2	90	1.1	0.179
E6		0.0	0.017

FIG. 20

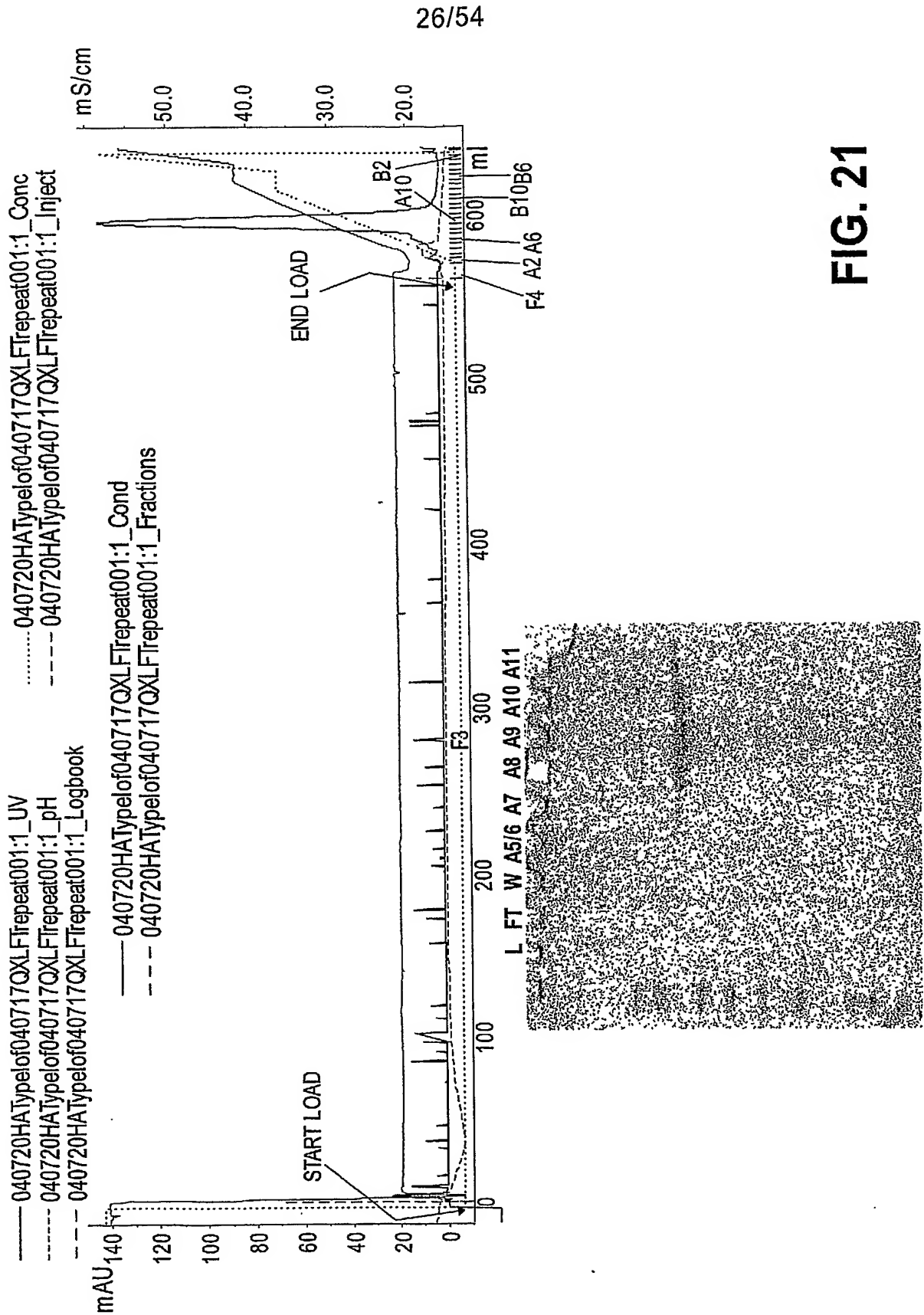


FIG. 21

27/54

Pooled A4-C4 and adjusted pH to 7.0 using 1mM HCl 670mL
load pH 7.0 16mS/cm

	Volume (mL)	Activity (U/L) pre- dialysis	Activity (U/L) post- dialysis	A280	A280/ 1.51 (mg/mL)	Activity (U)	Mass (mg)	Specific Activity (U/mg)
Load	670	9.2	NA	NA				
FT	670	0.0	NA	0.122				
Wash	9	2.9	NA	-0.013				
A5/6	6	1.1	NA	-0.005				
A7	3	0.1	19.3 (13mL)	0.180	0.119	0.25	1.55	0.16
A8	3	1.3						
A9	3	4.6						
A10	3	2.4						
A11	3	0.4						

FIG. 22

COMPARISON of MBP with MBP-SBD₃₉ tag in ST3Gal3 in Cyclodextrin column purification

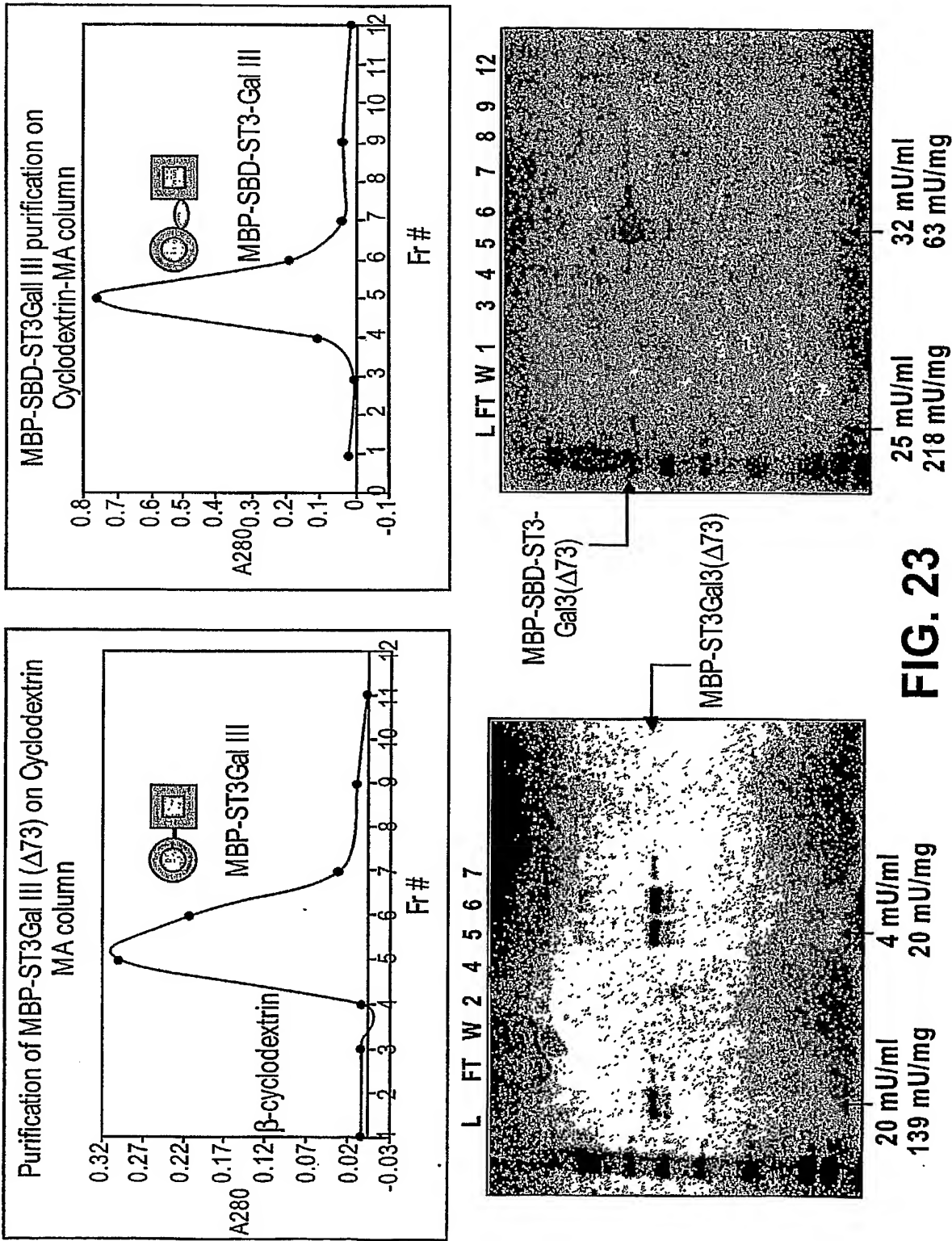


FIG. 23

29/54

MBP-pST3Gal1 fusion protein

MKIEGKLVINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIFWAHDF
 GGYAQSGLLAEITPDKAFQDKLYPFTWDVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALD
 KELKAKGKSALMFNLQEPYFTWPLIADGGYAFKYENGKYDIKDVGVNDNAGAKAGLFLVDLIKXKH
 MNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTPLPTFKGQPSKPFVGLSAGINAASP
 NKELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIAATMENAQKGEIMPNIPQMSAF
 WYAVRTAVINAASGRQTVDEALKDAQTNSSNNNNNNNNNNLHIEGRISEFGSELSENFKKLMKYPYR
 PCTCTRCIEEQRVSAWFDERFNRSMQPLLTAKNAHLEEDTYKWWLRLQREKQPNNLNDTRELQVVP
 GNVDPILLEKRLVSCRRCAVVGNSGNLKESYYGPQIDSHDFVLRMNKAPTEGFEADVGSKTTHHFVYPE
 SFRELAQEVSMILVPFKTTDLEWVISATTTGRISHTYVPVPAKIKVKKKELIYHPAFIKYVFDRWLQGH
 GRYPSTGILSVIFSLHICDEVLDYGFGADSKGNWHHYWENNPSAGAFRKTGVHDGDFESNVTTLASIN
 KIRIFKGR

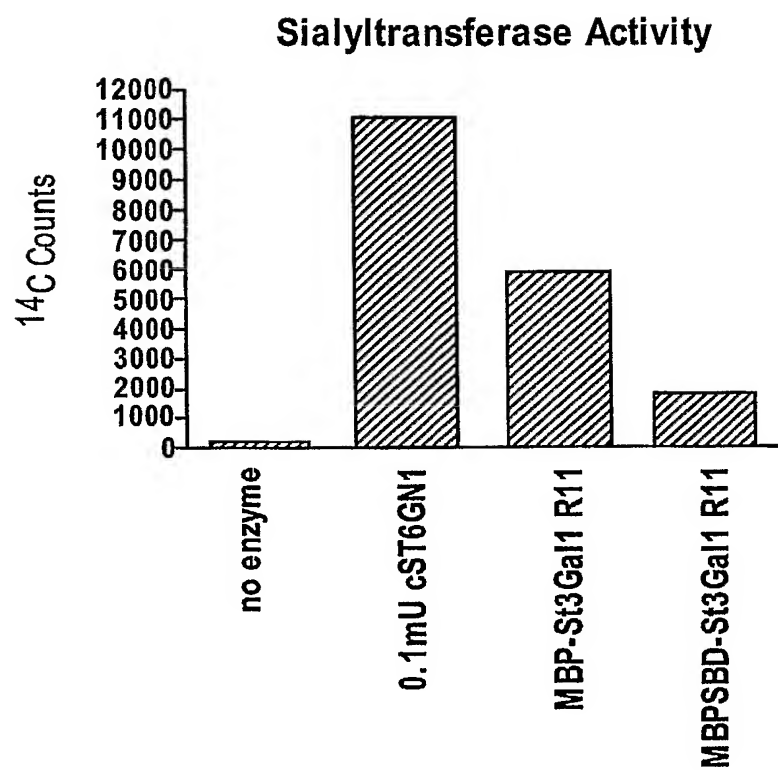
FIG. 24A

MBPSBD-pST3Gal1 fusion protein

MKIEGKLVINGDKGYNGLAEVGKKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIFWAHDF
 GGYAQSGLLAEITPDKAFQDKLYPFTWDVRYNGKLIAYPIAVEALSLIYNKDLLPNPPKTWEEIPALD
 KELKAKGKSALMFNLQEPYFTWPLIADGGYAFKYENGKYDIKDVGVNDNAGAKAGLFLVDLIKXKH
 MNADTDYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTPLPTFKGQPSKPFVGLSAGINAASP
 NKELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIAATMENAQKGEIMPNIPQMSAF
 WYAVRTAVINAASGRQTVDEALKDAQTNSSNNNNNNNNNNLHIEGRISEFGSIVATGGTTTATPTG
 SGSVTSTSKTTATASKTSTSTSTSTCTTPTAVAVTFDLTATTYGENIYLVGSISQLGDWETSDGIALSAD
 KYTSSDPLWYVTVTLPAGESEYKFIRESDDSVESDPNREYTPQACGTSTATVTDTWRGSELSEN
 FKLMKYPYRPCTCTRCIEEQRVSAWFDERFNRSMQPLLTAKNAHLEEDTYKWWLRLQREKQPNNLN
 DTIRELFQVVPGNVDPLLEKRLVSCRRCAVVGNSGNLKESYYGPQIDSHDFVLRMNKAPTEGFEADV
 SKTTHHFVYPESFRELAQEVSMILVPFKTTDLEWVISATTTGRISHTYVPVPAKIKVKKKELIYHPAFIK
 YVFDRWLQGHGRYPSTGILSVIFSLHICDEVLDYGFGADSKGNWHHYWENNPSAGAFRKTGVHDGDF
 ESNVTTLASINKIRIFKGR

FIG. 24B

30/54



Test pST3Gal1 for activity after Hampton refold

FIG. 25

31/54

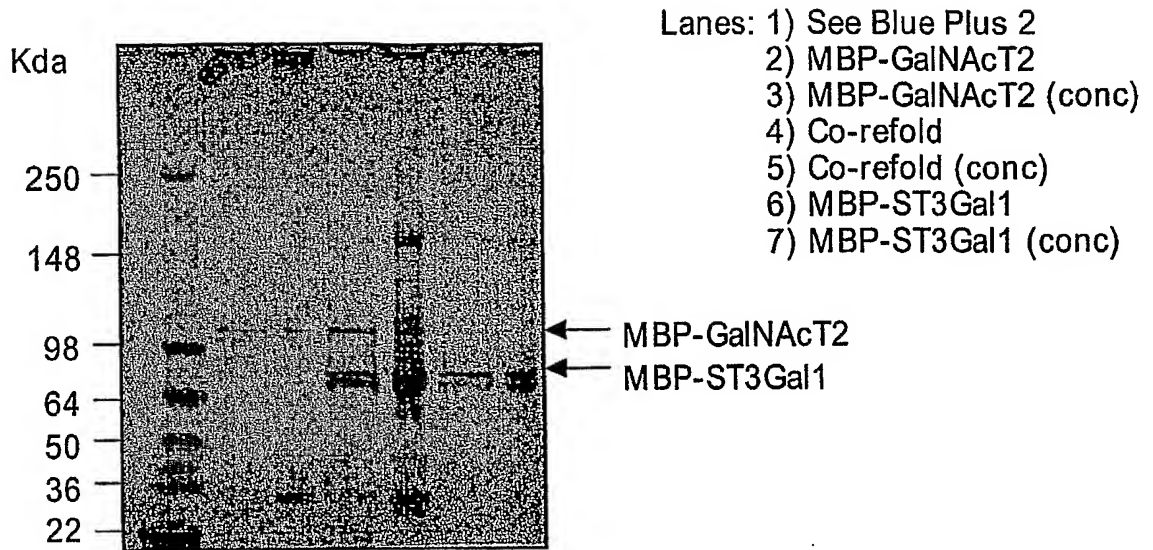
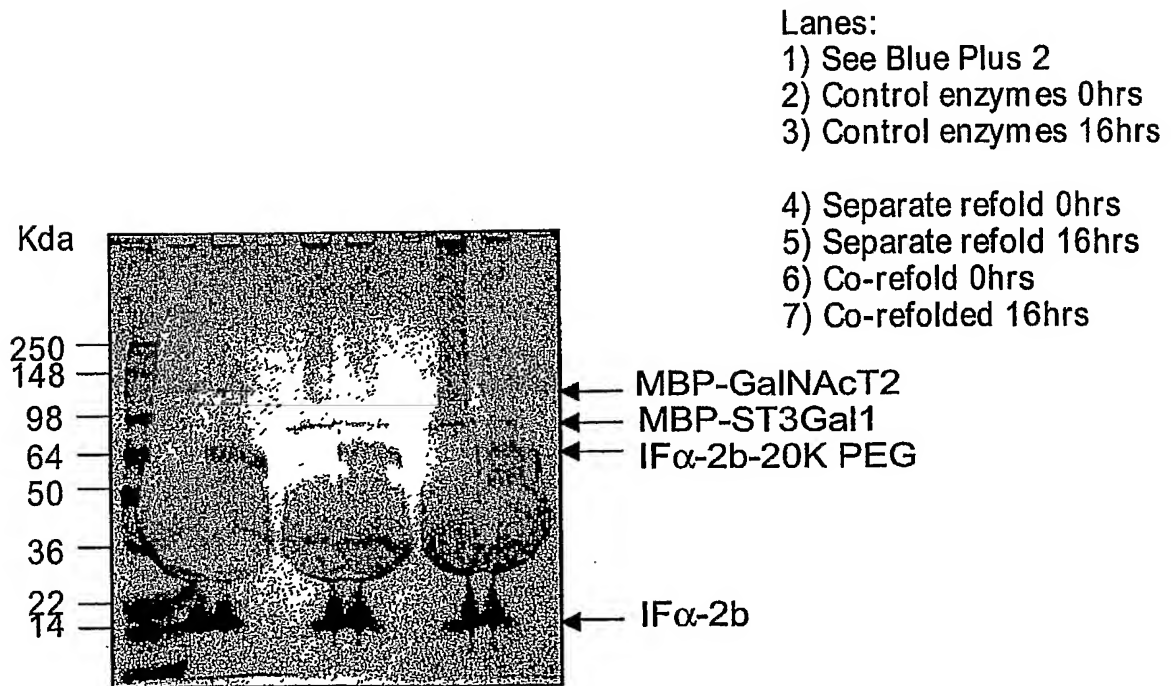
MKIEGKLVWINGDKGYNGLAEVGKKFEKDTGKVTVEHPDKLEEKFPQVAATGDGPDIFWAHDF
 GGYAQSGLLAEITPDKAFQDKLYPFTWDVRYNGKLIAYPIAVEALSLYNNKDLLPNPPKTWEEIPALD
 KELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGYDKIDVGVDNAGAKAGLFLVDLIKXKH
 MNADTDYSIAEAAFNKGETAMTINGPWAWNSNIDTSKVNYGVTVLPTFKGQPSKPFVGLSAGINAASP
 NKELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIAATMENAQKGEIMPNIPQMSAF
 WYAVRTAVINAASGRQTVDEALKDAQTNSSNNNNNNNNNNNNGIEGRSEFGSEHLLDKVPRITPGAL
 STRKTPMATGAVPAKKVVQATKSPASSPHPTTRRRQRRLKASEFKSEPRWDEEEYSLDMSSLQT
 NCSASVKIKASKSPWLQNFLPNITLFLDSGRFTQSEWNRLEHFAFPFGFMELNQSLVQKVVTFRFP
 PVRQQLLLASLPTGYSKCITCAVVGNGGILNDSRVGREIDSHDYVFRLSGAVIKGYEQDVGTRT
 SFYGFATFSLTQSILILGRRGFQHVPLGKDVRYLHFLGTRNRYEWLEAMFNLQTLAKTHLSWFR
 HRPQEAFRNALDLDRYLLHPDFELRYMKNRFLRSKTLDTAXWRIYRPTTGALLLLTALHLCCKV
 SAYGFITEGHERFSDHYDTSWKRLIFYINHDFLERMVWKRLLHDEGHWLYQRPQSDKAKN

FIG. 26A

MKIEGKLVWINGDKGYNGLAEVGKKFEKDTGKVTVEHPDKLEEKFPQVAATGDGPDIFWAHDF
 GGYAQSGLLAEITPDKAFQDKLYPFTWDVRYNGKLIAYPIAVEALSLYNNKDLLPNPPKTWEEIPALD
 KELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGYDKIDVGVDNAGAKAGLFLVDLIKXKH
 MNADTDYSIAEAAFNKGETAMTINGPWAWNSNIDTSKVNYGVTVLPTFKGQPSKPFVGLSAGINAASP
 NKELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIAATMENAQKGEIMPNIPQMSAF
 WYAVRTAVINAASGRQTVDEALKDAQTNSSNNNNNNNNNNNNGIEGRSEFGSKEPQTKPSRHQKTE
 NIKERSLQSLAKPKSQAPTRARITTYAEPVPENNALNTQTQPKAHTTGDGRGKEANQAPPEEQDK
 VPHTAQRAAWKSPEKEKTMVNTLSRPGQDAGMASGRTEAQSWSQDTKTTQCGNGGQTRKLT
 SRTVSEKHQGAATTAKTLIPKSQHRMLAPTGAVSTRQKGVTTAVIPPEKKPQATPPAPFQ
 SPTTQRNORLKAANFKSEPRWDFEEKYSFEIGGLQTTCPDSVKIKASKSLWLQKFLPNLTLFLDS
 RHFNQSEWDRLEHFAFPFGFMELNYSLVQKVVTFRPPVPOQQLLASLPAGSLRCITCAVVGNG
 GILNNSHMGQEIDSHDYVFRLSGALIKGYEQDVGTRTSFYGFTAFSLTQSLILGNRGFKNVPLGK
 DVRYLHFLGTRDYEWLEALLMNQTVMSKNLFWFRHRPQEAFREALHMDRYLLLHPDFLRYM
 KNRFRLRSKTLDDGAHWRIYRPTTGALLLTALQLCDQVSAYGFTEGHERFSDHYDTSWKRLIFY
 INHDFKLEREVWKRLLHDEGHIRLYQRP GGTAKAKN

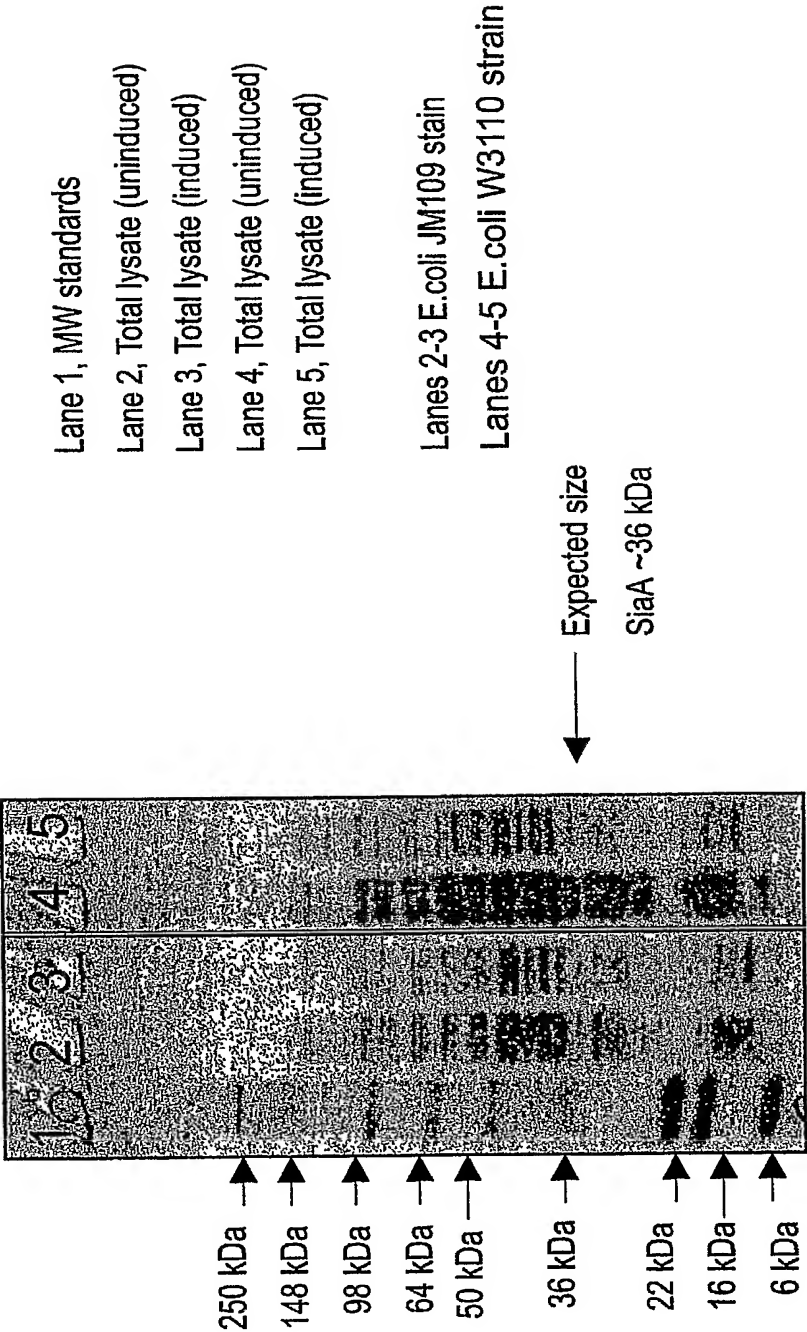
FIG. 26B

32/54

**FIG. 27A****FIG. 27B**

33/54

Induced SiaA expression in E.coli



There is no obvious inducible band at the expected
Mass of 36 kDa for the native SiaA protein.

FIG. 28

34/54

Induced SiaA/MBP expression in E.coli

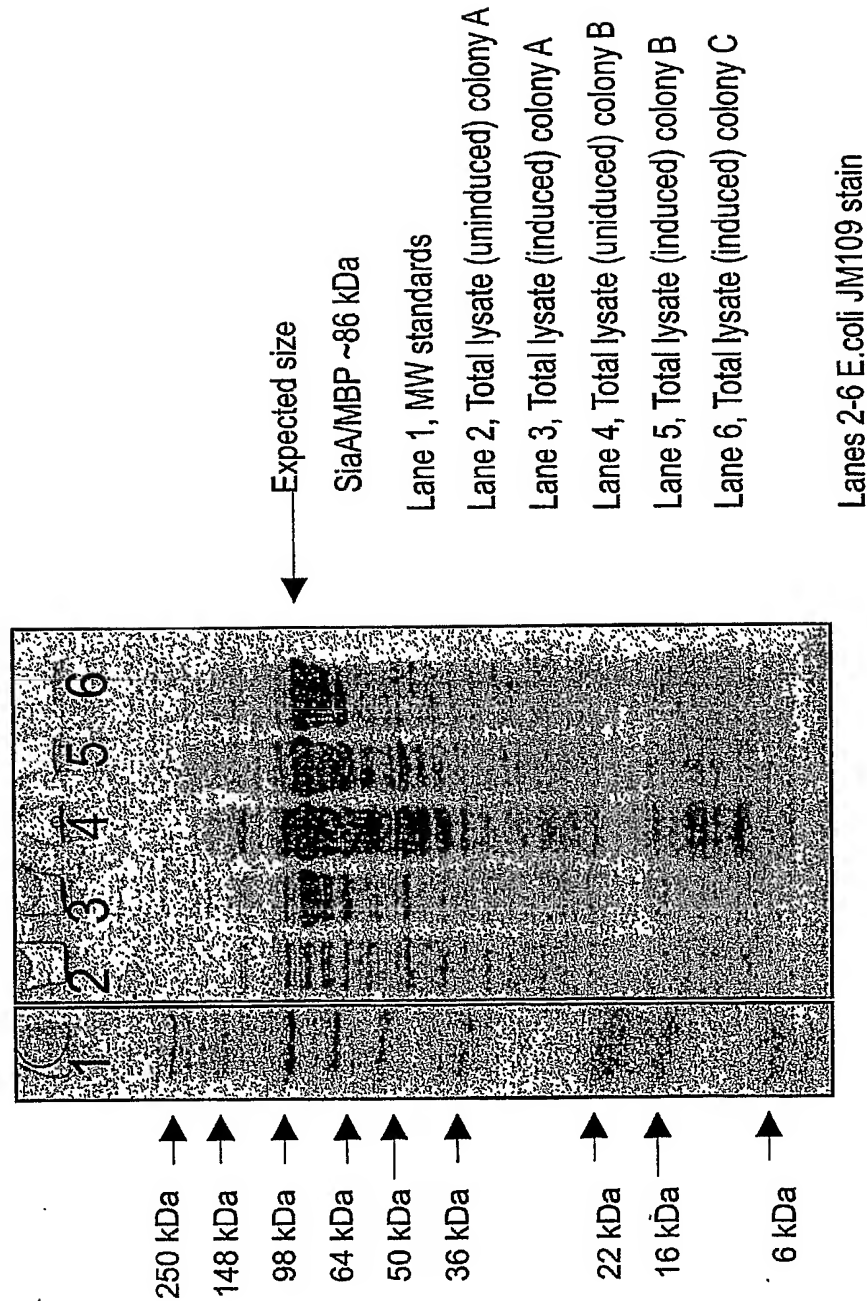


FIG. 29

High level production of SiaA/MBP even in absence of IPTG induction.
Compare to figure X where SiaA production is not obvious. The presence of the fusion partner (MBP) drives high levels of expression.

35/54

```

1  mkfrepllgg saampgaslq racrllvavc alhlgvtlv ylagrdlrrl pqlvgvhppl
61  qgsshgaaai gqpsgelrlr gvaoppplqn sskprsraps nldayshpgp gpgpgsnlts
121  apvpstttrs ltacpeespl lvgpmliefn ipvdlklieq qnpkvklggr ytpmdcisp
181  kvaiiilfrn rqehlkywly ylhpmvqrqq ldygiyvinq agesmfnrak llnvgfkeal
241  kdydyncfvf sdvdlipmnd hntyrctfsqp rhisvamdkf gfsipyvqyf ggvsalskqq
301  flsingfpnn ywgwggeddd iynrlafrgm svsrpnavig krmirhsrd kknepnqrf
361  driahtketm lsdglnslty mvlevqrypl ytkitvdigt ps

```

FIG. 30

Bovine (b) GalT1 (β 1,4GalT1) constructs

UDP-galactose β -N-acetylglucosaminide β 4-galactosyltransferase (EC 2.4.1.38)

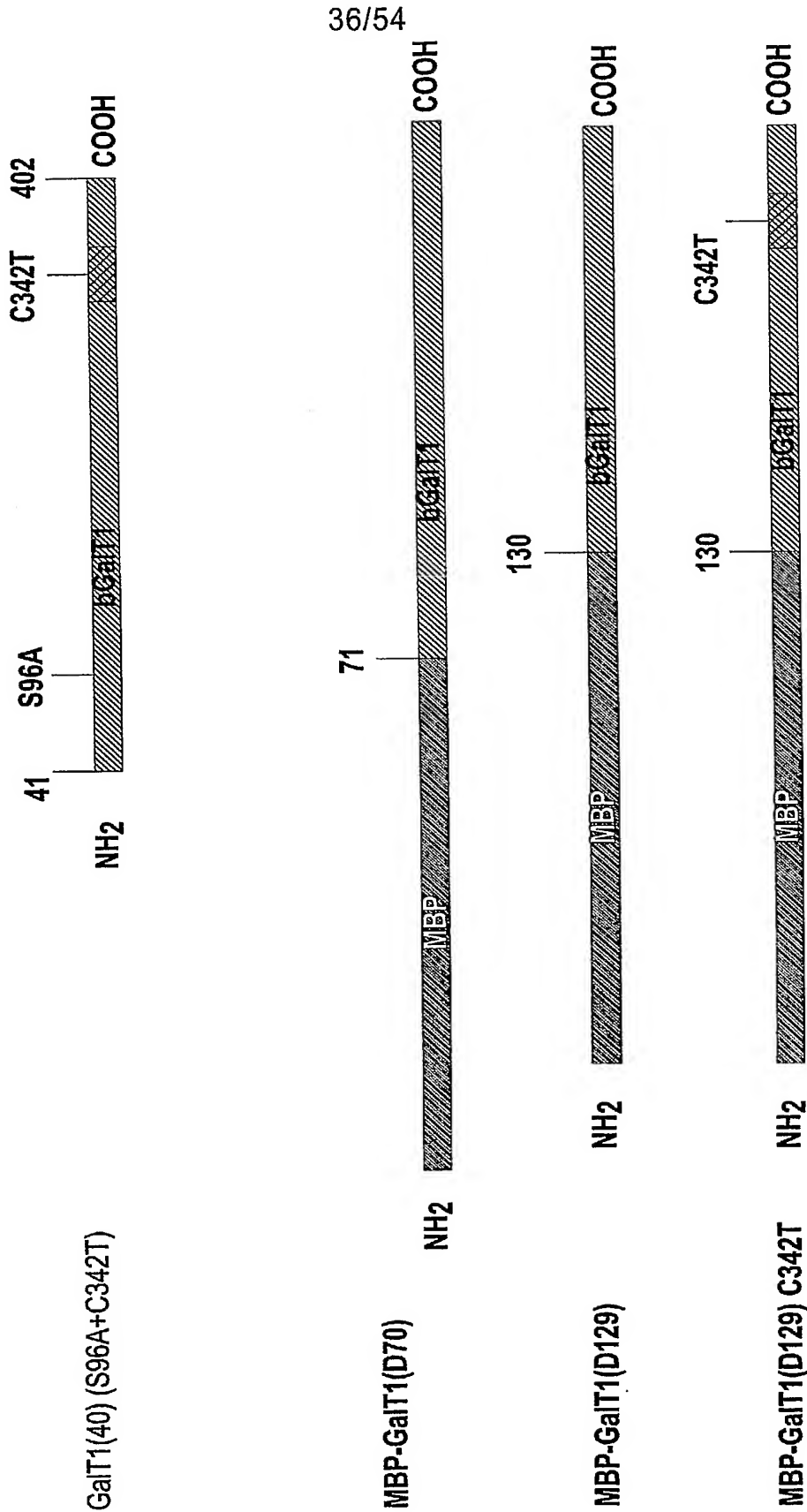
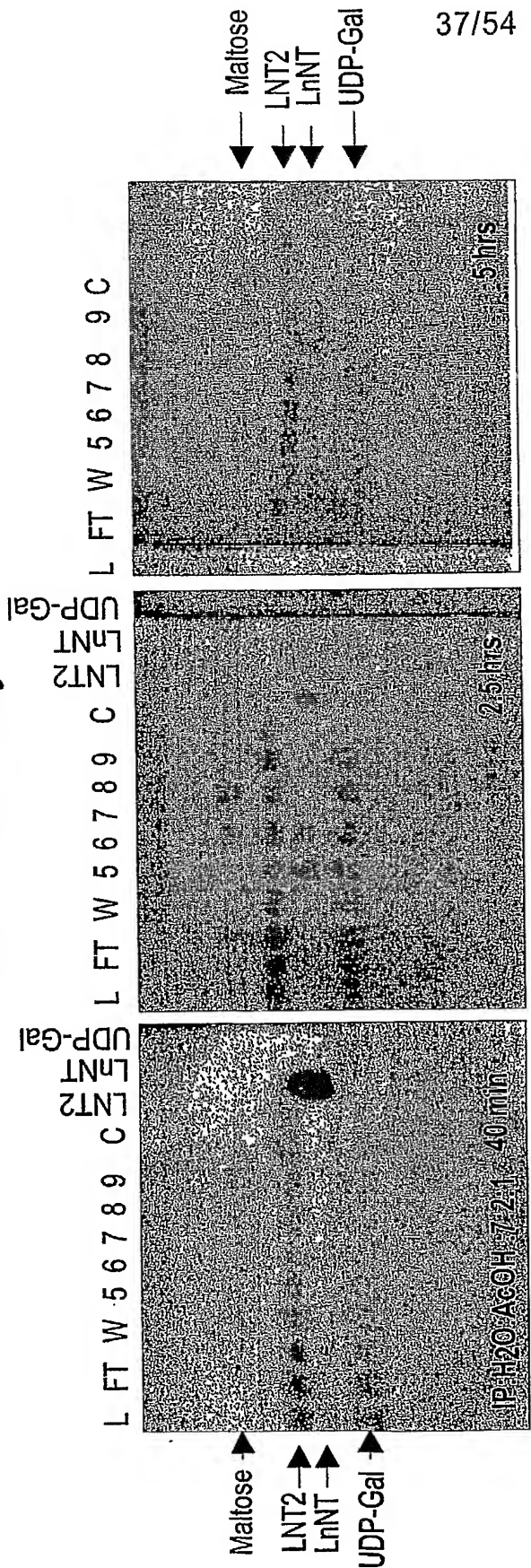


FIG. 31

GalT1 TLC assay



37/54

LnNT reactions with

- L: Loaded sample (refolded, dialyzed MBP-GalT1(70) onto amylose column)
- FT: Amylose column Flow trough
- W: Amylose column wash
- 5: Maltose Eluted Fr # 5
- 6: Maltose Eluted Fr # 6
- 7: Maltose Eluted Fr # 7
- 8: Maltose Eluted Fr # 8
- 9: Maltose Eluted Fr # 9
- C: Control with water

FIG. 32

38/54

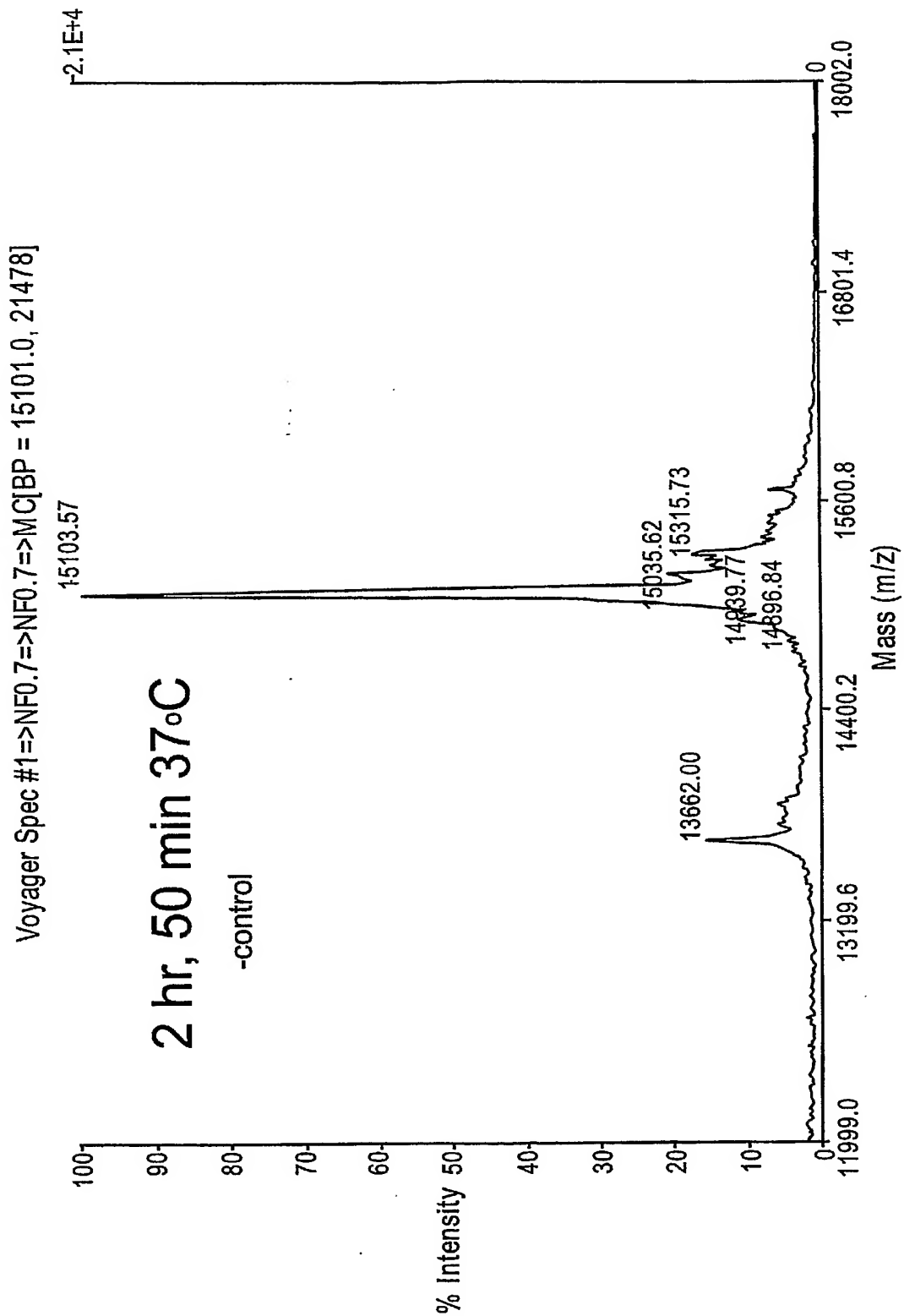


FIG. 33

39/54

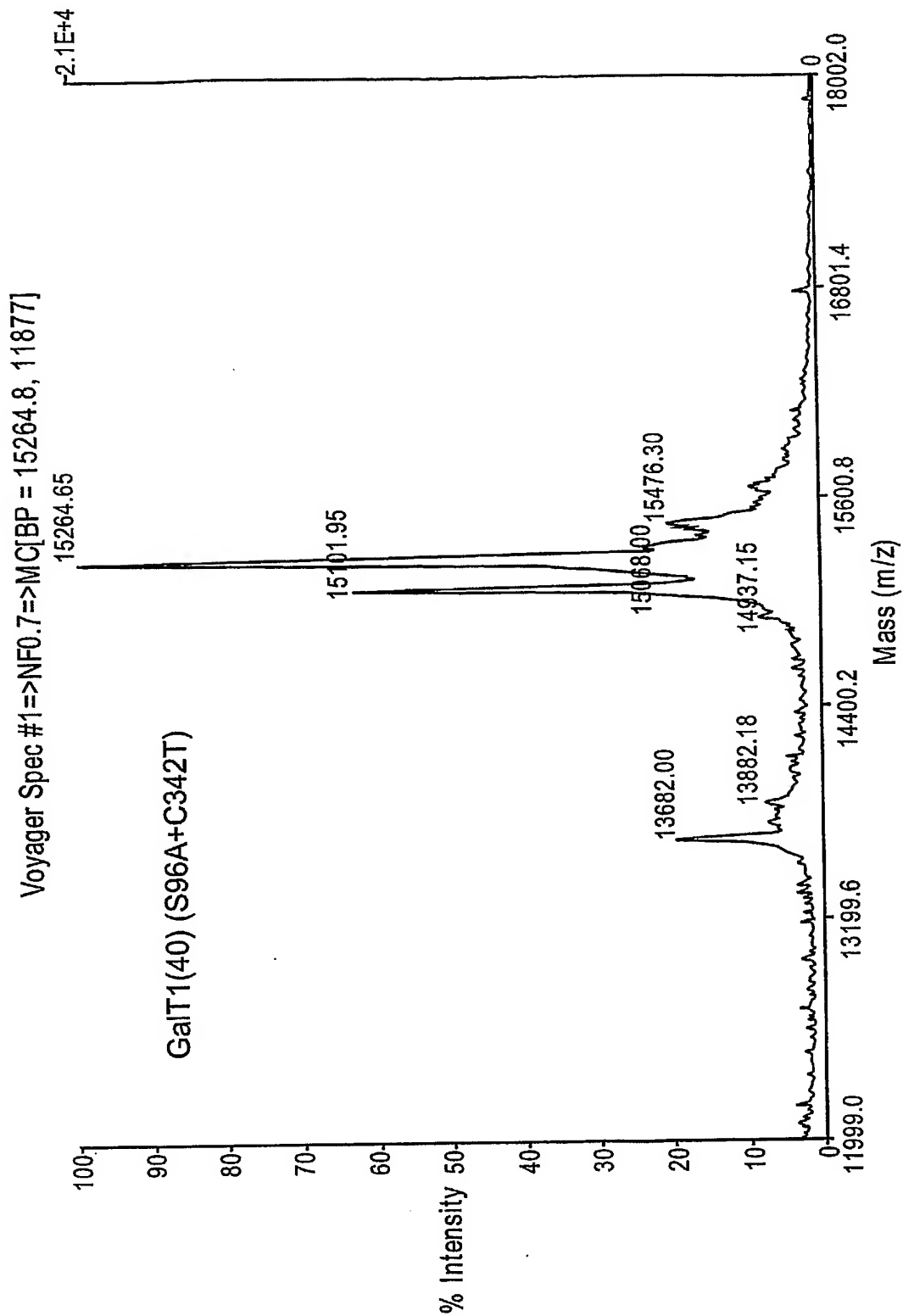


FIG. 33 (CONT.)

40/54

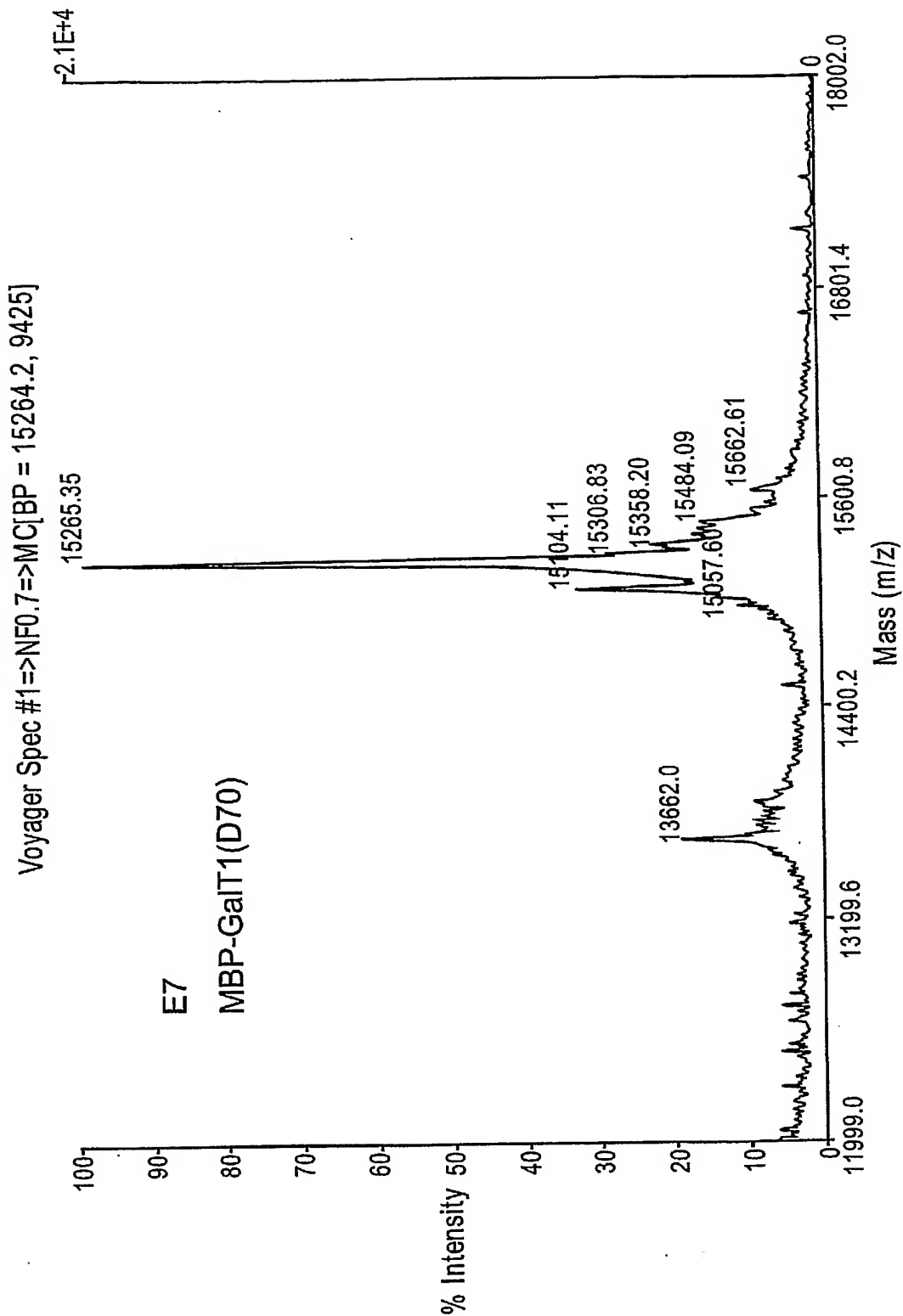


FIG. 33 (CONT.)

41/54

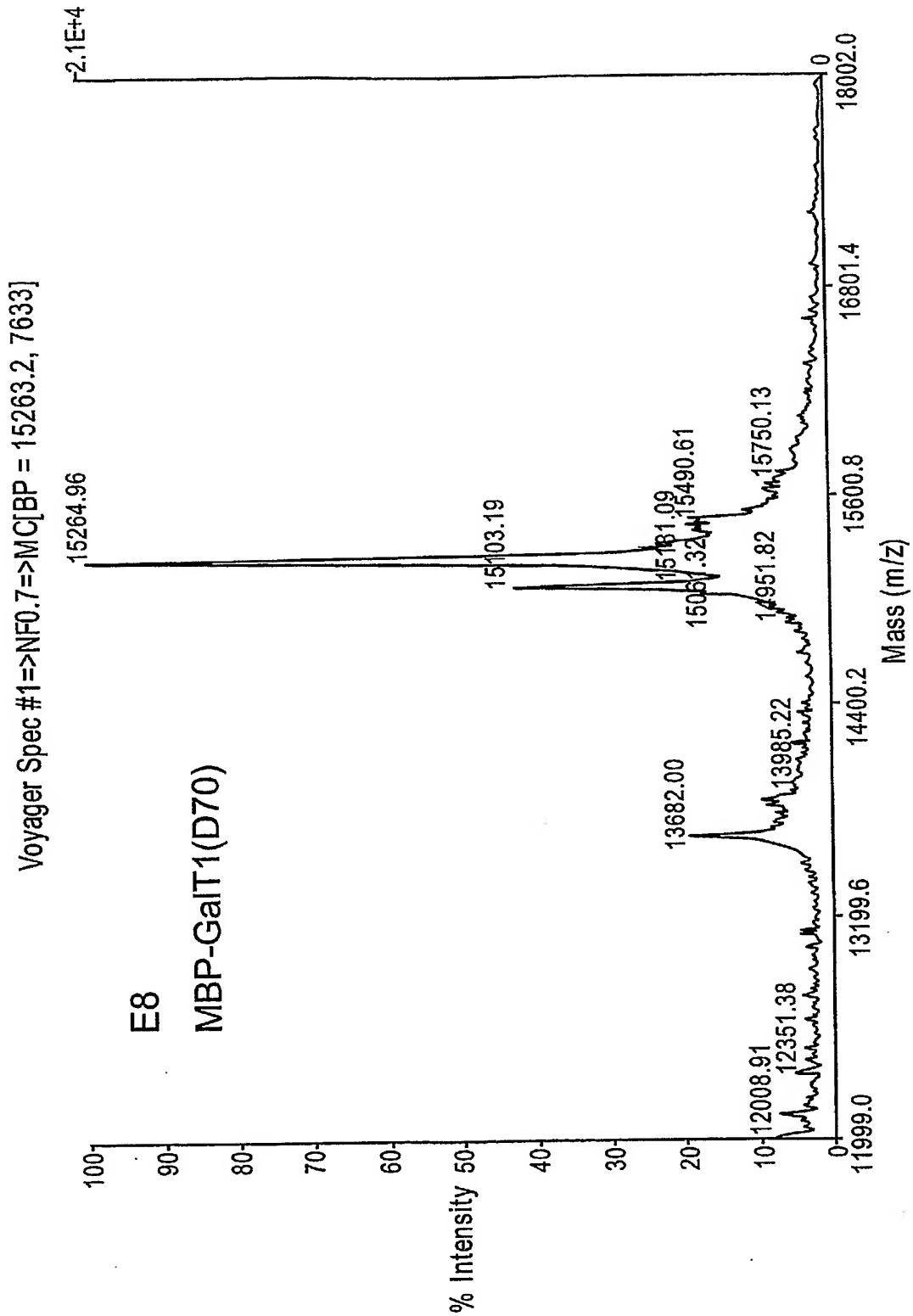


FIG. 33 (CONT.)

42/54

Kinetics of RNase B modelling with GalT1 6.5.03

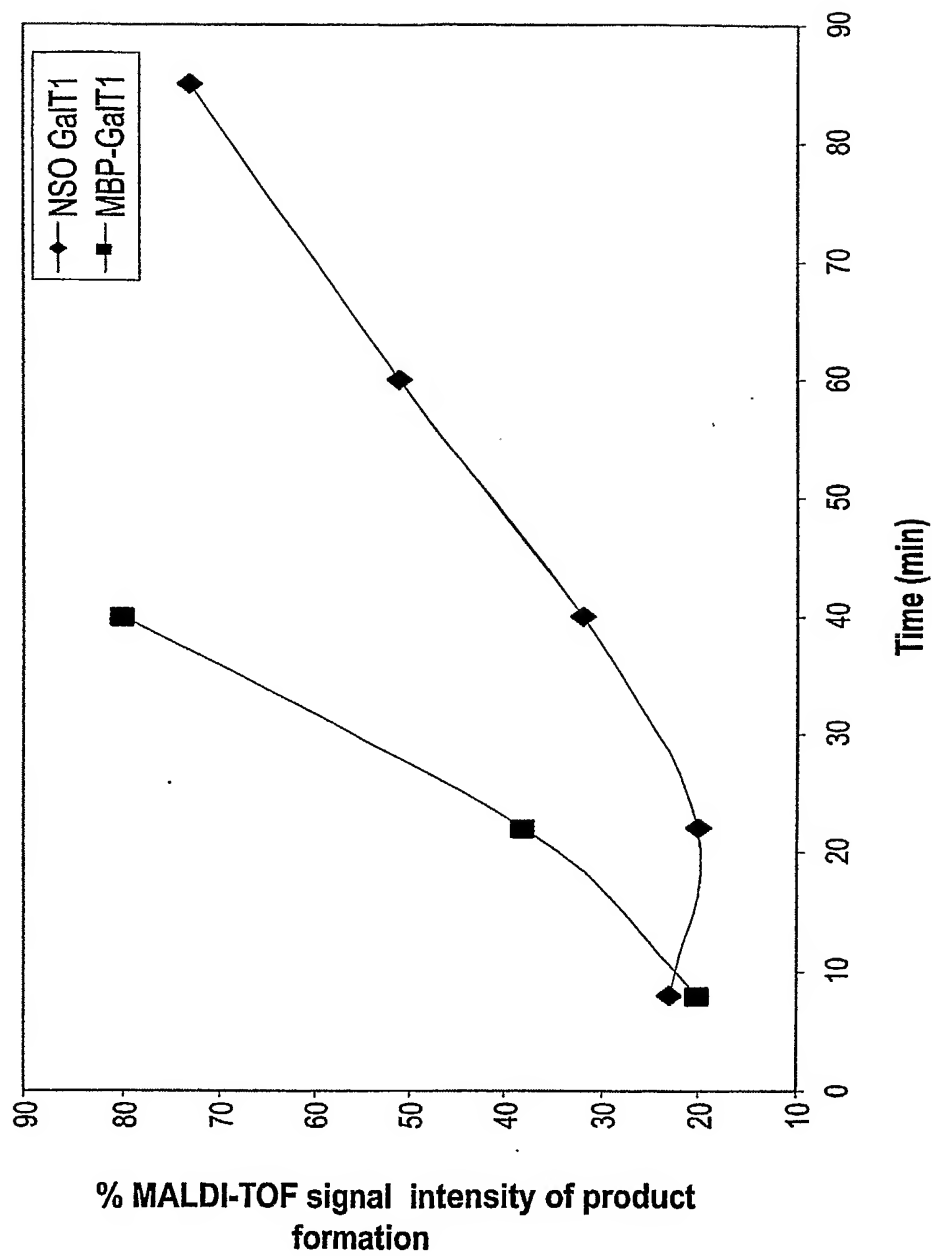
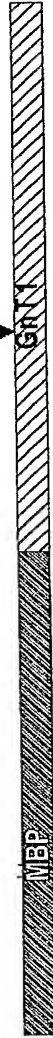


FIG. 34

GnT1 constructs

MBP-GnT1(Δ 35)

C121

MBP-GnT1(Δ 103)

43/54

1 mlkksaglv lwgailfvaw nallllffwt rpapgrppsv saldgdpasl trevirlaqd
 61 aevelerqrg llqqigdals sqgrvrptaa ppagrvpvt papavipilv iacdrstvir
 121 clckllhyrp saelfpiivs qdcgheetaq aiasygsavt hirqpdlsi avppdhrkfq
 181 gyykiahryr walgqvfrqf rfpaavvved dlevapdffe yfratypllk adpslwcvs
 241 wndngkeqmv dasrpellyr tdffpglgwl llaelwaele pkwpkafwdd wmrpeqrqg
 301 racirpeisr tmtfgrkgvs hggffdqhlk fiklngqfvh ftqldlsylq reaydrdfla
 361 rvygapqlqv ekvrtndrke lgevrvqytg rdsfkafaka lgvmdllksg vpragyrgiv
 421 tfqfgrrvrh lappptwegy dpswn

FIG. 35

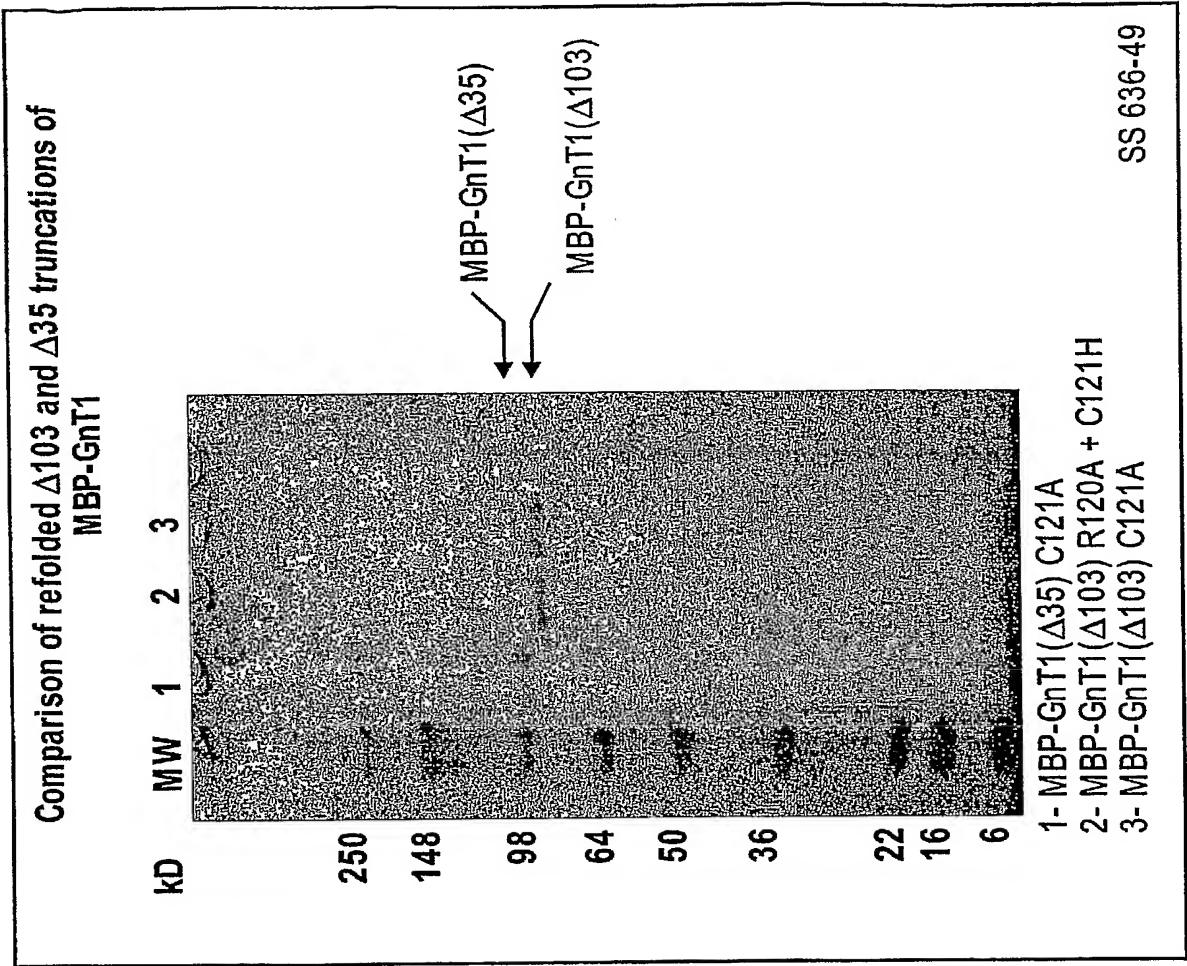
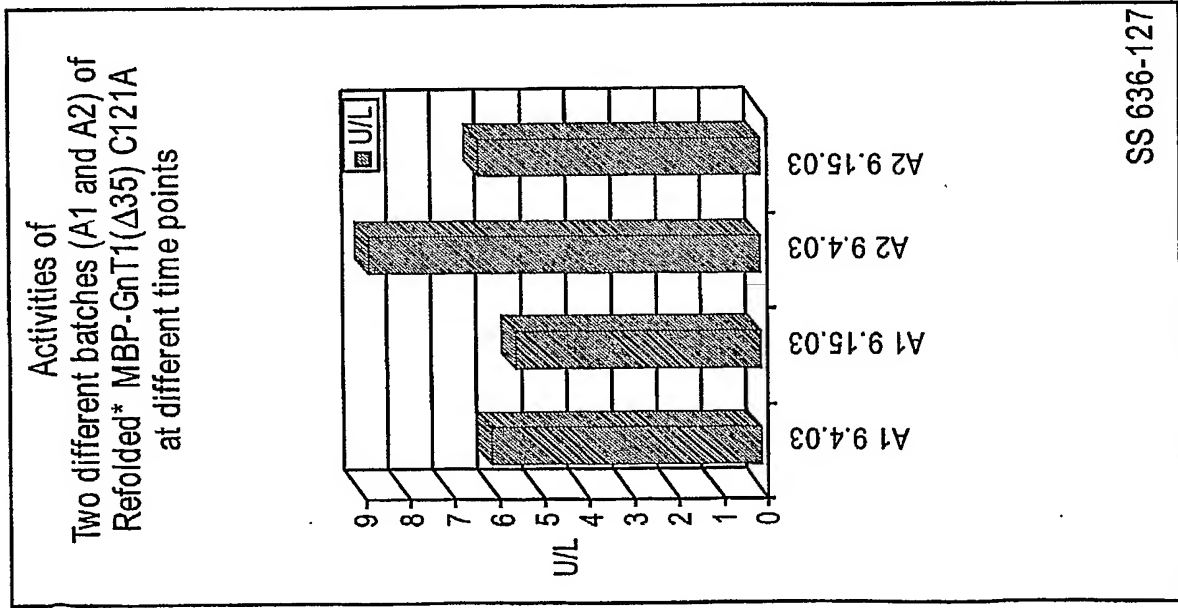


FIG. 36



45/54

1 mapmrkkstl klltllvlfi fltsfflnys htvvttawfp kqmvielsen fkklmkypyr
61 pctctrciee qrvsawfder fnrsmqpllt aknahleedt ykwrlrlqre kqpnnlnndti
121 relfqvvp gn vdpallekrlv scrrcavvgn sgnlkesyvg pqidshdfvl rmnkaptgef
181 eadvgsktth hfvypesfre laqevsmilv pfktttlewv isattgtis ht yvpvpaki
241 kvkkekiliiy hpafikyvfd rwlqghgryp stgilsvifs lhicdevdly gfgadskgnw
301 hhywennpsa gafrktgvhd gdfesnvtti lasinkirif kgr

FIG. 37

46/54

Human ST6GalNAcI

MRSCLWRCRHL SQGVQWSLL LAVLVFFLFALPSFIKEPQTKPSRHQRTENIKERSLQS
 LAKPKSQAPTRARRTTIYAEPVPENNALNTQTQPKAHTTGDRGKEANQAPPEEQDK
 VPHTAQRAAWKSPEKEKTMVNTLSPRGQDAGMASGRTEAQSWKSQDTKTTQNGG
 GQTRKLTASRTVSEKHQGAATTAKTLIPKSQHRMLAPTGA VSTRTRQKGVTTAVIP
 PKEKKPQATPPPAPFQSPTTQRNQR LKAANFKSEPRWDFEEKYSFEIGGLQTTCPDSV
 KIKASKSLWLQKLFLPNLTLFLDSRHFNQSEWDRLEHFAPPF GFMELNYSLVQKVVT
 RFPPVPQQQLLLASLPAGSLRCITCAVVGNGGILNNSHMGQEIDSHDYVFR LSGALIK
 GYEQDVGTRTSFYGFTAFSLTQSLLILGNRGFKNVPLGKDVRYLHFLEGRDYEWLE
 ALLMNQTVMSKNLFWFRHRPQEA FREALHMDRYLLLHPDFLRYMKNRFLRSKTL D
 GAHWRIYRPTTGALLLLTALQLCDQVSAYGFITEGHERFSDHY YDTSWKRLIFYINH
 DFKLEREVWKRLHDEGIIRLYQRPGPGTAKAKN

FIG. 38A**Chicken ST6GalNAcI**

MGFLIRRLPKDSRIFRWLLILT VFSFIITSFSALFGMEKSIFRQLKIYQSI AHMLQVDTQ
 DQQGSNYSANGRISKVGLERDIAWLELNTAVSTPSGEGKEEQKKT VKPVAKVEEAK
 EKVTVKPFPEVMGITNTTASTASVVERTKEKT TARPVPGVGEADGKR TTIALPSMKE
 DKEKATVKPSFGMKVAHANSTSKDKPKAEPPASVKAIRPVTQAATVTEKKKLRAA
 DFKTEPQWDFDDEYILDSSSPVSTCSESVRAKAAKSDWLRDLFLPNITL FIDKSYFNV
 SEWDRLEHFAPPYGFME LNYSLVVEEVM SRLPPNPHQQLLANSSSNVSTCISCAVVG
 NGGILNNSGMGQEIDSHDYVFRVSGAVIKGYEKDVGT KTSFYGFTAYSLVSSLQNLG
 HKGFKKIPQGKHIRYIHFLEAVRDYEWL KALLLDKDIRKGFLNYYGRRPRERFDEDF
 TMNKYLVAHPDFLRYLKNRFLKSKNLQKPYWRLYRPTTGALLLLTALHLCDRV SAY
 GYITEGHQKYS DHYYDK EWKRLVFYVNHDFNLEKQVWKRLHDENIMKLYQRS

FIG. 38B

Mouse ST6GalNAcI protein beginning at residue 32 of the native mouse protein
 DPRAKDSRCQFIWKNDASAQENQQKAEPQVPIMTLSPRVHNKESTSVSSKDLKKQER
 EAVQGEQAEGKEKRKLETIRPAPENPQSKAEPAAKTPVSEHLDKLPRTPGALSTRKTP
 MATGAVPAKKKVQATKSPASSPHPTTRRRQRLKASEFKSEPRWDFEE EYSLDMSSL
 QTNCASV KIKASKSPWLQNIFLPNITLFLDSGRFTQSEWNRLEHFAPPF GFMELNQSL
 VQKVVT RFPPVRQQQLLLASLPTGYSKCITCAVVGNGGILNDSRVGREIDSHDYVFR
 LSGAVIKGYEQDVGTRTSFYGFTAFSLTQSILILGRRGFQHVPLGKDVRYLHFLEGR
 NYEWLEAMFLNQTLAKTHLSWFRHRPQEA FRNALDLDRYLLLHPDFLRYMKNRFL
 RSKTLDTAHWRIYRPTTGALLLLTALHLCDKVSAYGFITEGHQRFSDHY YDTSWKRL
 IFYINHDFRLERMVWKRLHDEGIWLYQR PQSDKAKN

FIG. 38C

47/54

Hum (h)ST6GalNAcI truncations

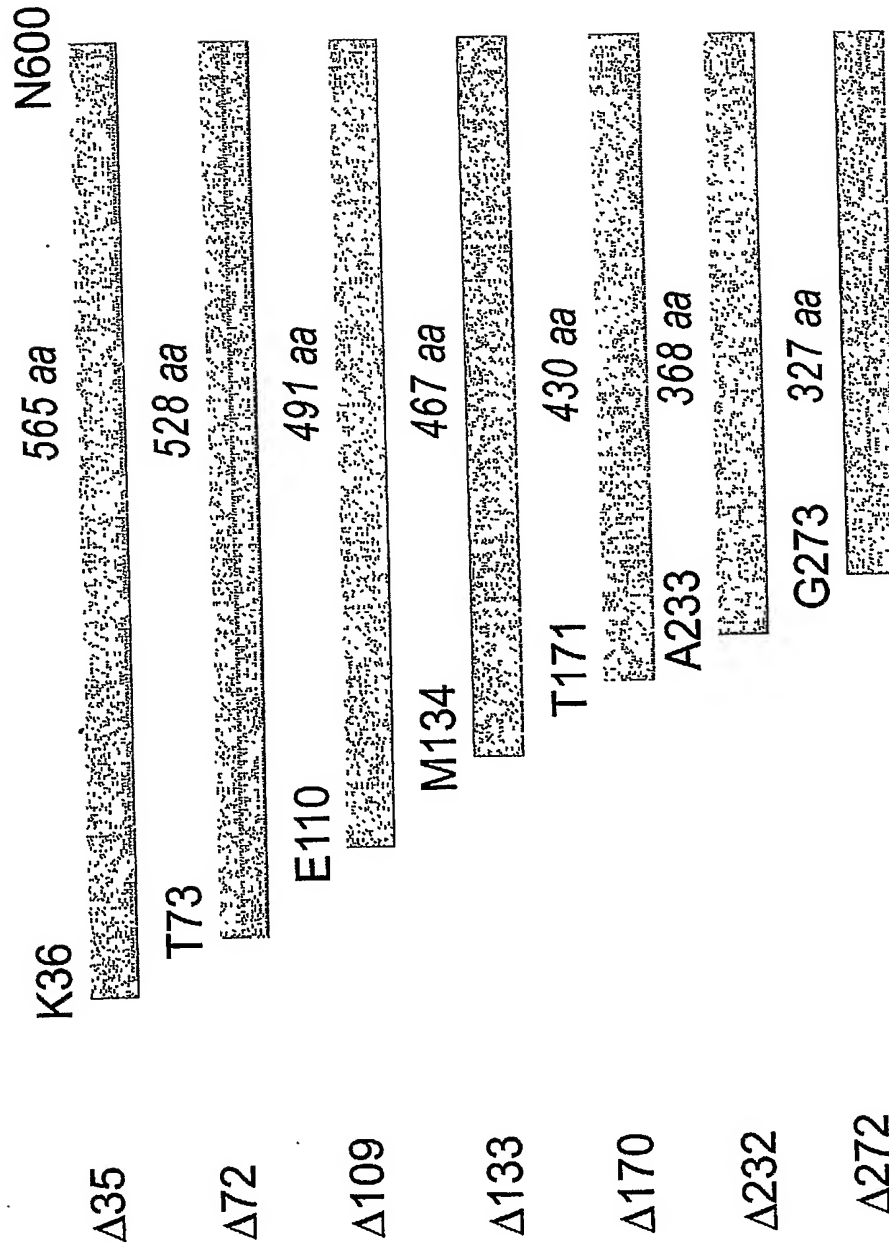


FIG. 39

48/54

MBP-hST6GalNAc constructs

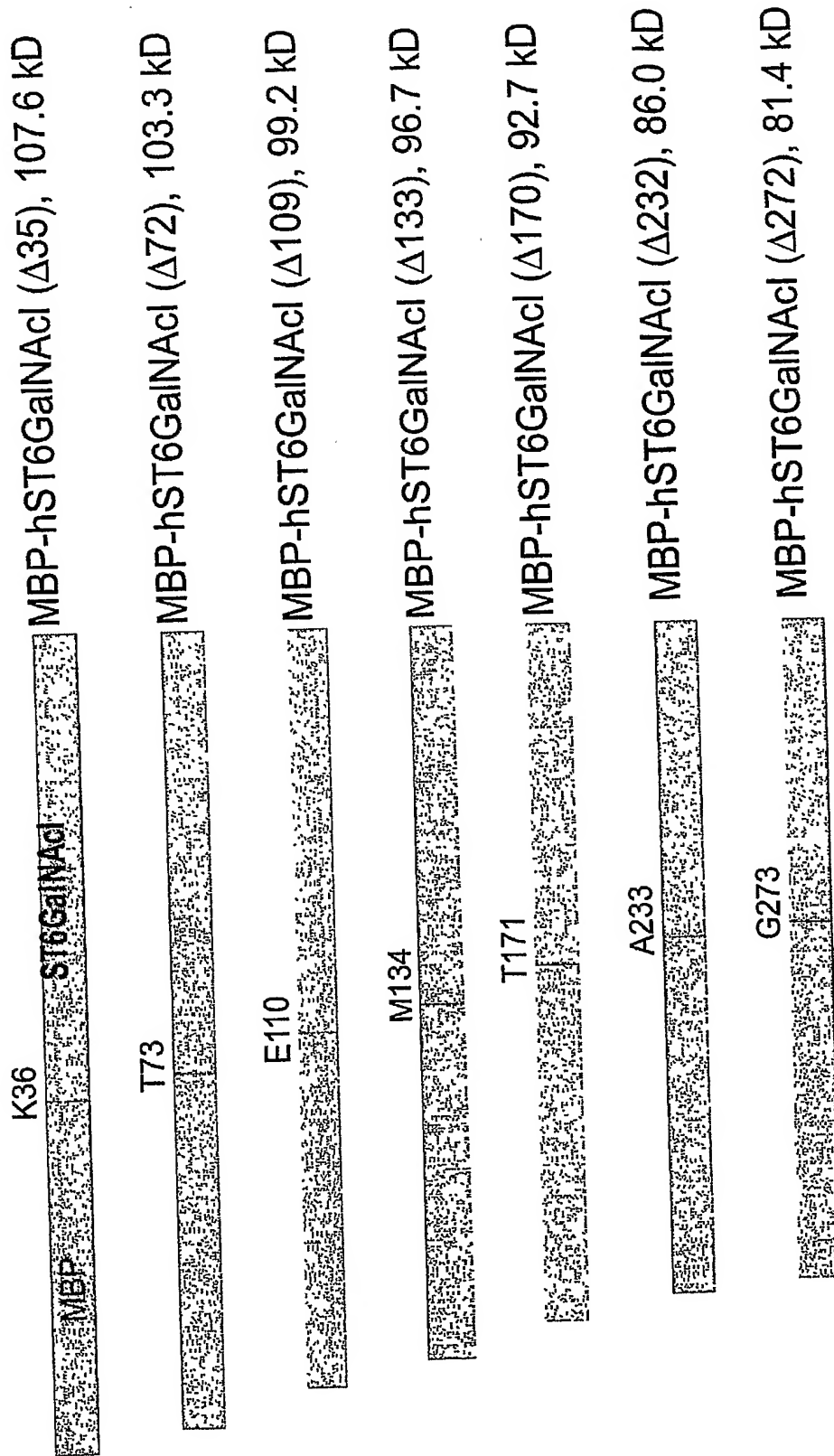


FIG. 40

49/54

MASKSWLNFLTFLCGSAIGFLLCSQLFSILLGEKVDTQPNVLHNDPHARHSDDNGQN
HLEGQMNFNADSSQHKDENTDIAENLYQKVRILCWVMTGPQNLEKKAKHV KATW
AQRCNKVLFMSSEENKDFPAVGLKTKEGRDQLYWKTIKAFQYVHEHYLEDADWFL
KADDDTYVILDNLRWLLSKYDPEEPIYFGRRFKPYVKQGYMSGGAGYVLSKEALKR
FVDAFKTDKCTHSSSIEDLALGRCMEIMNVEAGDSRDTIGKETFHFPVPEHHLIKGYL
PRTFWYWNYNYYPPVEGPGCCSDLAVSFHYVDSTTMYELEYLVYHLRPYGYLYRY
QPTLPERILKEISQANKNEDTKVKLGNP

FIG. 41

Sequence	Size	Range	Mode
SC1GALT1.AMI	342	1-	Normal
SPTS122.AMI	342	1-	Normal
		10	20 30 40 50
SC1GALT1.AMI	1	EFPYDGRH	DMMMSGPEQ
SPTS122.AMI	1	EFPYDGRH	DMMMSGPEQ
		60	70 80 90 100
SC1GALT1.AMI	51	EVRVLCWIMT	NKLI FMSSAK
SPTS122.AMI	51	EVRVLCWIMT	NKLI FMSSAK
		110	120 130 140 150
SC1GALT1.AMI	101	VGEGRNNLWG	LKADDDTYTI
SPTS122.AMI	101	VGEGRNNLWG	LKADDDTYTI
		160	170 180 190 200
SC1GALT1.AMI	151	YSPETPVYFG	SREAVRRFVV
SPTS122.AMI	151	YSPETPVYFG	SREAVRRFVV
		210	220 230 240 250
SC1GALT1.AMI	201	SDNSGAEDVE	HHLIPSHTDK
SPTS122.AMI	201	SDNSGAEDVE	HHLIPSHTDK
		260	270 280 290 300
SC1GALT1.AMI	251	KFWYWQYIFY	PNQMYVL DYL
SPTS122.AMI	251	KFWYWQYIFY	PNQMYVL DYL
		310	320 330 340 350
SC1GALT1.AMI	301	NTPDALPNKL	VSKRS A ET KT
SPTS122.AMI	301	NTPDALPNKL	VSKRS T ET KT

FIG. 42

51/54

Yersinia MBP

MKIEEGKLV	WINGDKGYNG	LAEVGKKFEK	DTGIKVTIEH	PKLEEEKFPQ
VAATGDGPD	IFWAHDRFGG	YAQSGLLAE	TPSKAFQEK	FPFTWDVAVR
NGKLIGYPV	VEALSIIYNK	DLVKEAPKTW	EEIPALDKTL	RANGKSAIMW
NLQEPYFTWP	VIAADGGYAF	KEFENGVDYDAK	NVGVNNAGAQ	AGLQFIVDLV
KNKHINADTD	YSIAEAAFNK	GETAMTINGP	WAWSNIDKSK	INYGVTLLPT
FHGQPSKPFV	GVLTAGINAA	SPNKELATEF	LENYLITDQG	LAEVNKKDKPL
GAVALKSFQE	QLAKDPRIAA	TMDNATNGEI	MPNIPQMAAF	WYATRSVAVLN
AITGRQTVEA	ALNDAATRIT	K		

FIG. 43A*E. coli* MBP

MKIEEGKLV	WINGDKGYNG	LAEVGKKFEK	DTGIKVTVEH	PKLEEEKFPQ
VAATGDGPD	IFWAHDRFGG	YAQSGLLAEI	TPDKAFQDKL	YPFTWDVAVR
NGKLIAYPIA	VEALSIIYNK	DLNPNPPKTW	EEIPALDKEL	KAKGKSALMF
NLQEPYFTWP	LIAADGGYAF	KYENGKYDIK	DVGVDNAGAK	AGLTFLVDLI
KNKHMNADTD	YSIAEAAFNK	GETAMTINGP	WAWSNIDTSK	VNYGVTVLPT
FKGQPSKPFV	GVLSAGINAA	SPNKELAKEF	LENYLLTDEG	LEAVNKKDKPL
GAVALKSYEE	ELAKDPRIAA	TMENAOQKEI	MPNIPQMSAF	WYAVRTAVIN
AASGRQTVDE	ALKDAQTNS			

FIG. 43B*Pyrococcus furiosus* MBP

MKIEEGKVVI	WHAMQPNELE	VFQSLAE EYM	ALSPEVEIVE	EQKPNLEDAL
KAAIPTGQGP	DLFIWAHDWI	GKFAEAGLLE	PIDEYVTEDL	LNEFAPMAQD
AMQYKGHYA	LPFAAETVAI	IYNKEMVSEP	PKTFDEMKA	MEKYYDPANE
KYGIAPWPIA	YFISAIQAQAF	GGYYFDDKTE	QPGLDKPETI	EGFKFFFTTEI
WPYMAPTGDY	NTQQSIFLEG	RAPMMVNGPW	SINDVKKAGI	NFGVVPLPPI
IKDGKEYWPR	PYGGVKLIYF	AAGIKNKDAA	WKFAKWLTS	EESIKTALAE
LGYIPVLTKV	LDDPEIKNDP	VIYGFQAVQ	HAYLMPKSPK	MSAVWGGVDG
AINEILQDPQ	NADIEGILKK	YQQEILNNMQ	G	

FIG. 43C*Thermococcus litoralis* MBP

MKIEEGKIVF	AVGGAPNEIE	YWKGVIAEFE	KKYPGVTVEL	KRQATDTEQR
RLDLVNALRG	KSSDPDVFLM	DVAWLGGQFIA	SGWLEPLDDY	VQKDNVDLSV
FFQSVINLAD	KQGGKLYALP	VYIDAGLLYY	RKDLLEKYGY	SKPPETWQEL
HEMAQKIQSG	ERETNPNFWG	FVWQGGQYEG	LVCDFVEYVY	SNGGSLGEFK
DGKWVPTLNK	PENVEALQFM	VDLIHKYKIS	PPNTYTEMTE	EPVRLMFQQG
NAAFERNWPY	AWGLHNADDS	PVKGKVGVP	LPHFPGHKSA	ATLGGWHIGI
SKYSDNKALA	WEFVKFVESY	SVQKGFAMNL	GWNPGRVVDY	DDPAVVS KSP
HLKELRAVFE	NAVPRPIVPY	YPQLSEIIQK	YVNSALAGKI	SPQEALDKAQ
KEAEELVKQY	SK			

FIG. 43D*Thermatoga maritima* MBP

MKIEQTKLTI	WSSEKQVDIL	QKLGEFEKAK	YGIPVEVQYV	DEFSIKSKFL
TAAPQGGQAD	IIVGAHDWVG	ELAVNGLIEP	IPNFSDLKNF	YDTALKAFSY
GGKLYGVPIA	MEAVALIYNK	DYVDSVPKTM	DELIEKAKQI	DEEYGGGEVRG
FIYDVANFYF	SAPFILGYGG	YVFKETPQGL	DVTDIGLANE	GAVKGAKLIK
RMIDEGVLTP	GDNYGTMDSM	FKEGLAAMII	NGLWAIKSYK	DAGINYGVAP
IPELEPGVPA	KPFVGVQGF	INAKSPNKVI	AMEFLT NFIA	RKETMYKIYL
ADPRLPARKD	VLELVKDNPD	VVAFTQSASM	GTPMPNVPEM	APVWSAMGDA
LSIIINGQAS	VEDALKEAVD	KIKAQIEK		

FIG. 43E

SUBSTITUTE SHEET (RULE 26)

52/54

Vibrio cholerae MBP

MKIEEGQITI WINGDKGYNG LAEVGKKFEA DTGIKVTVAH PDALQDKFPQ
TAATGDGPD I VFWAHDREFGG YAEAGLLVEI KPSAKIQEGI VDFAWDAVKY
NGKIIGYPIA VESLSLIYNK DLVPNPPKSW EEVAELDAKL KKEGKSAIMW
NLKEPYFTWP LMAADGGYAF KYGVDGYDVK DAGINNKGVK DAMNEVKGLV
DKGVISPDMD YSVSESANQ GNTAMTINGP WSWGNIKSG INYGVTTLPK
FNGQASKPFV GVLTAGISTA SPNKDLAVEF IENYLLTNDG LRMVNNDKPL
GAVALNSFQR ELDADARIAA TMDNAMNGEI MPNIPQMNAF WSSAKNAIIN
IVDGRQTVDA ALADAQMT KP

FIG. 43F

53/54

	1	50
HSGALNAT1.pep	(1)MRKFAYCKVVLATSLIWVLLDMFILLLYFS-----	ECNKC
HSGALNAT2.pep	(1)MRRRS--RMLLCFAFLWVLGIAYMYSGGGSALAGGAGGGAGRKEDWNEI	
Consensus	(1)MRK A KMLL A IWVL F L	D N
	51	100
HSGALNAT1.pep	(35)DEKKERGLPAGDVLEPVQKPHEGP-G-----	EMGKPVVIPKEDQEKMKEM
HSGALNAT2.pep	(49)DPIKKKDLHHSNGEKAQSMETLPPGKVRWPDFNQEAYVGGTMVRSGQDP	
Consensus	(51)D K K L E Q P G D I	D
	101	150
HSGALNAT1.pep	(79)FKINQENLMASEMIALNRSPLDVRLEGCKTKVYPDNLPTTSVVIVFHNEA	
HSGALNAT2.pep	(99)YARNKFNQVESDKLRMDRAIPDTRHDQCQRKQWRVDLPATSVVITFHNEA	
Consensus	(101)F N FN M SD I L RAIPD R D C K W LP TSVVI FHNEA	
	151	200
HSGALNAT1.pep	(129)WSTLLRTVHSVINSRPHMIEEIVLVDDASERDFLKRPLESYVKKLKVPV	
HSGALNAT2.pep	(149)RSALLRTVSVLKKSPPHLIKEIILVDDYSN-----	DPEDGALLGKIEKV
Consensus	(151) S LLRTV SVI KSP HLI EIILVDD S	P D L V
	201	250
HSGALNAT1.pep	(179)HVIRMEQRSGLIRARLKGA AVSKGOVITFLDAHCECTVGWLEPLIARIKH	
HSGALNAT2.pep	(194)RVLNRDRREGLMRSRVRGADAAQAKVLTFLDSHCECNEHWLEPLIERVAE	
Consensus	(201) VIR D R GLIRARLKGA A A VITFLDAHCEC	WLEPLL RI
	251	300
HSGALNAT1.pep	(229)DRRTVVCPIIDVISDDTFEYMAGSMTYGGFNWKLNFRWYPVPQREMDRR	
HSGALNAT2.pep	(244)DRTRVVSPIIDVINMDNFQYVGASADLKGGFDWNLVFKWDYMTPEQRRSR	
Consensus	(251)DR VV PIIDVI D F YMAAS	GGF W L FKW M R
	301	350
HSGALNAT1.pep	(279)KGDRTL PVRTPTIMAGGLFSIDRDYFQEIGTYDAGMDIWGGENLEISFRIW	
HSGALNAT2.pep	(294)QGNPVAPIKTPTMIAGGLFVMDKFYFEELGKYDMMMDVWGGENLEISFRVW	
Consensus	(301) G PIKTP IAGGLF IDK YF EIG YD MDIWGGENLEISFRIW	
	351	400
HSGALNAT1.pep	(329)QCGGTLEIVTCSHVGHVFRKATPYTFPGGTGQIINKNNRRLAEVWMDEFK	
HSGALNAT2.pep	(344)QCGGSLEIIPCSRVGHVFRKQHPYTFPGSGTVFARNTRRAAEVWMDEYK	
Consensus	(351)QCGGSLEII CS VGHVFRK PYTFPGSG I KN RR AEWWMDEFK	
	401	450
HSGALNAT1.pep	(379)NFFYIISPGVTKVDYGDISSRVGLRHKLQCKPFSWYLENIYPDSQIPRHY	
HSGALNAT2.pep	(394)NFFYAAVPSARNVPYGNIQSRLELRKKLSCKPFKWYLENVYPELRVPDHO	
Consensus	(401)NFFY P V YG I SRL LR KL CKPF WYLENIYPD	IP H
	451	500
HSGALNAT1.pep	(429)FSLGEIRNVEITNQCLDNMARKENEKVGIFNCHGMGGNQVFSYTANKEIRT	
HSGALNAT2.pep	(444)DIAFGALQOGTN-CLDTLGHFADGVVGVYECHNAGGNQEWALTREKSVKH	
Consensus	(451) N TN CLD LA VGIF CH GGNQ FA T K IK	
	501	550
HSGALNAT1.pep	(479)DDLCLDVSKLN--GPVTMLKCHHLKGNQLWEYDPVKLTLOHVNSNQCLDK	
HSGALNAT2.pep	(493)MDLCLTVVDRAPGSLIKLQGCRENDSRQKWEQIEGNSKL RHVGSNLCLDS	
Consensus	(501) DLCL V I L C Q WE L HV SN CLD	
	551	584
HSGALNAT1.pep	(527)ATEEDSQVPSIRDCNGSRSQOWLLRNVTLP EIF-	
HSGALNAT2.pep	(543)RTAK-SGGLSVEVCGPALSQOWKFTLNLOQ----	
Consensus	(551) T S SI C A SQOW	

FIG. 44

SUBSTITUTE SHEET (RULE 26)

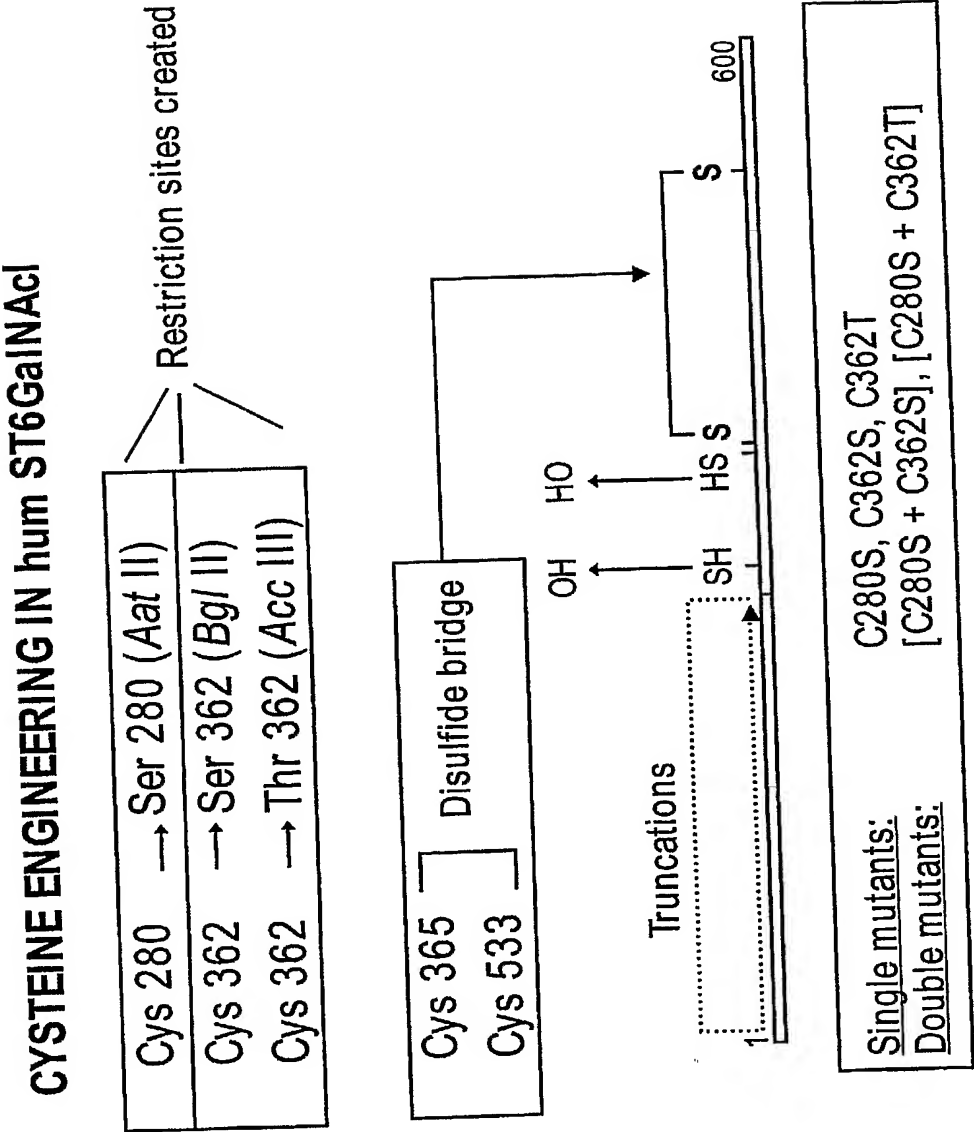


FIG. 45